Postal Regulatory Commission
Submitted 6/8/2015 3:52:13 PM
Filing ID: 92506
Accepted 6/8/2015

BEFORE THE
POSTAL REGULATORY COMMISSION

Periodic Reporting (Proposal Thirteen)	:	**Docket No. RM2015-7**

UNITED PARCEL SERVICE COMMENTS ATTACHING SUPPLEMENTAL REPORT RELATED TO PROPOSAL THIRTEEN

(June 8, 2015)

In Order No. 2455, the Commission requested that United Parcel Service, Inc. ("UPS") submit a supplemental report detailing the results of additional analyses performed after UPS accessed a non-public Crosswalk File for Form 3999 data. UPS hereby submits the Supplemental Report from Dr. Kevin Neels of the Brattle Group, describing this further work and its results, which is attached hereto as Exhibit A. Below, UPS provides an overview of this work and responds briefly to points made by the Postal Service in its Reply Comments dated May 13, 2015.

I. OVERVIEW

UPS is participating in this docket because it believes that Postal Service costing practices for this large and important cost segment can and should be improved. For years, cost attribution in this segment has relied upon flawed special studies conducted well over a decade ago. The model based on those studies was not fully litigated and was accepted at least in part because the Commission had no viable alternative. This model has consistently yielded very small attribution to competitive products. Although

competitive products contributed 21.1% of the Postal Service's total revenue in FY2013, the Postal Service costing methodology assigned just 6.9% of total City Carrier Street Time costs to competitive products, and less than 9.2% of attributable costs to competitive products.[1] This means that – despite the huge growth in competitive products delivered by the Postal Service and the large investments the Postal Service is making in delivering competitive products – market dominant mailers were required to bear *over 93%* of the costs of City Carrier Street Time.

Proposal Thirteen would move the needle only slightly: under Proposal Thirteen, market dominant mailers would still bear about 90% of the costs of this segment. This is because Proposal Thirteen continues to employ costing approaches that have not meaningfully changed from the legacy approach. Instead of letting the actual data speak for itself, Proposal Thirteen relies upon *a priori* assumptions about how narrowly and artificially defined cost categories behave, and it also relies heavily on data collected from costly special field studies that have inherent limitations.

UPS acknowledges the time and resources the Postal Service has spent on developing Proposal Thirteen. UPS is aware of the complexities involved in developing sound costing practices. But there is substantial room for improvement in the costing model for this segment, and the Commission and the public benefit when private parties spend time and resources analyzing the available data and proposing alternative approaches for the Commission's consideration. Accurate cost attribution to competitive products is critical to meet Congress' mandates in the Postal Accountability

[1] *See* FY2013 Cost and Revenue Analysis; FY2013 Cost Segments and Components Worksheet (available at https://about.usps.com/who-we-are/financials/); calculation assumes competitive products are assigned an "appropriate share" of 5.5% of institutional costs, in addition to the reported competitive attributable costs.

and Enhancement Act ("PAEA"). Yet, far too often after PAEA, the Commission is presented with no alternatives to the models proposed by the Postal Service.

As noted in our initial comments, UPS's objectives in this docket are limited to improving the calculation of *marginal* cost in this cost segment. UPS strongly believes that attribution based on the marginal cost of the last unit (*i.e.* attributing what the Postal Service misleadingly calls "volume variable" costs), while failing to attribute other variable costs, does not comport with PAEA. UPS is preparing a separate petition that will address these and other problems that permeate Postal Service cost attribution practices for competitive products beyond this particular cost segment.

As discussed in his Supplemental Report, to improve the marginal cost calculation, Dr. Neels has conducted a series of analyses using the non-public Crosswalk File obtained by UPS. This Crosswalk File allowed Dr. Neels to analyze the Postal Service's Form 3999 dataset, which was extracted from the Postal Service's operational data systems and contains route evaluations for 140,457 city carrier routes across more than 10,000 ZIP Codes nationwide. Dr. Neels' work demonstrates that costing models for City Carrier Street Time can and should eschew the Postal Service's heavy reliance on assumptions and costly special studies and move toward an approach grounded in actual data. For all of the Postal Service's criticisms regarding Dr. Neels' preliminary work, neither the Postal Service nor its expert, Professor Bradley, seriously dispute that this is a worthy objective.[2]

[2] Dr. Neels responds to certain of these criticisms in his Supplemental Report. Overall, UPS notes that many of the criticisms concern work that Dr. Neels acknowledged was preliminary and based on the very limited data he had at the time or that he presented solely to highlight limitations inherent in Proposal Thirteen.

The Commission should not adopt a model – such as Proposal Thirteen – that *assumes* that parcel delivery does not affect so-called "regular" delivery time or that certain cost components are wholly fixed or how various cost activities interact, when those assumptions are unnecessary. Similarly, the Commission should not again adopt for use in this and coming years a model that is heavily reliant on data that must be collected from expensive special field studies, when other options are available. As Dr. Neels explains: "A strategy that is based upon the results of costly special studies runs a chronic risk of being forced to rely on analyses that are under-powered, over-simplified and out of date." Supplemental Neels Report at 2.

Replacing assumptions and special studies with a holistic analysis of operational data will yield more accurate and more useful results. And it is important to make this change now, in order to begin accounting for the growing importance of parcels to the Postal Service's business. While the Postal Service has experienced declines in overall mail volume in recent years, parcel volumes have grown sharply and continue to rise. These changes underscore the need for a new approach that yields results that better reflect ongoing changes in the mail stream.

Dr. Neels has developed a new approach for the Commission to consider. With the benefit of being able to use the more robust dataset from the Form 3999 database, Dr. Neels' recommended approach has evolved from his preliminary work in connection with his initial report, where he was severely constrained by the limited data available to him at that time. Based on his analysis of which model best fits the additional data, Dr. Neels now utilizes the flexible functional form model used by the Postal Service in Proposal Thirteen, but with some important modifications. Specifically, Dr. Neels' model

takes a more holistic (and realistic) approach, eliminating the artificial division of mail delivery time into “parcel” and “regular” delivery time. Dr. Neels’ model also removes other *a priori* assumptions, including that one cost pool is wholly fixed and that the drivers of so-called “allied” activities are identical to the drivers of the corresponding direct street activities.

As Dr. Neels explains, this approach yields strong results. His model – the “National Form 3999 Model” – considers all city carrier street time in a comprehensive and consistent fashion and dispenses with untested and increasingly questionable assumptions about cost structure and cost causation. It is based on a much larger and richer set of data than Proposal Thirteen, and it provides results that are correspondingly more precise and reliable. It increases the attribution to competitive products in a manner that reflects the increasing role of these products. And because it is based on operational data, the analysis can be updated next year, or whenever desired, at a modest cost. For these reasons, UPS respectfully submits that the Commission should adopt Dr. Neels’ National Form 3999 Model.

This model admittedly has limitations driven by deficiencies in the Postal Service’s data: because of the inadequate collection of parcel-related data and data on customer collection volume, the National Form 3999 model must for now impute volume measures for a few postal products. Although Dr. Neels has developed imputation approaches that are sufficient for present purposes, better data collection practices would eliminate the need for these imputations and yield better results in the future. Accordingly, the Commission should also direct the Postal Service to improve its collection of parcel-related data in the Form 3999 dataset.

II. UPS'S CRITICISMS OF PROPOSAL THIRTEEN

UPS writes briefly to explain why the Postal Service's responses to UPS's comments about Proposal Thirteen are unpersuasive.

A. The Postal Service's Exclusion Of Parcels From The Regular Delivery Equation Is Inappropriate.

By excluding parcels from the "regular" delivery equation, Proposal Thirteen embeds an assumption in the model that there is no relationship between parcel and "regular" delivery time. While Proposal Thirteen purports to measure parcel-delivery time, the only costs it attributes to parcels come from the cost pools for In-Receptacle Package Delivery and Deviation Delivery, as well as a portion of the indirect "allied" cost pools. *See* CCST Study at 19. This means that Proposal Thirteen assumes that the presence of parcels does not impact the cost pool called "Regular Delivery," which supposedly represents the time associated with the delivery of non-parcel postal products. That assumption is not supported by the data, and it defies common sense.

As Dr. Neels explains, it is not reasonable to assume, today, that the presence of parcels has no effect on the cost pool the Postal Service calls "Regular Delivery Time." Over the course of the day, a letter carrier handles a number of different mail streams, including parcels, and those streams are intermingled. It is implausible that special studies involving mail carriers holding timers can accurately measure the discrete amount of time spent dealing with any one of those mail streams. Dr. Neels analogizes this to trying to unscramble eggs to figure out how much time was spent cooking the

egg whites. Neither the Postal Service nor Dr. Bradley provide any compelling reason to conclude this can be done accurately.[3]

The Postal Service argues that its approach is consistent with an "assumption" that was accepted by the Commission a decade ago. *See* Postal Service Reply Comments at 20 (May 13, 2015) ("This is not a new assumption."); Bradley Analysis of Neels Report at 10-11 (May 13, 2015) ("Bradley Report"). But there is no reason to rely upon old assumptions that are testable by data using modern econometric methods. A lot has changed in the Postal Service's business over the past decade, including the significant growth in parcel volumes.

The Postal Service also argues that "in no instance, does UPS present any evidence to support these assertions." Postal Service Reply Comments at 19. But Dr. Neels *did* provide statistical evidence testing this assumption: in his initial report, Dr. Neels presented results that demonstrate the impact of including a variable for parcels in the Proposal Thirteen model for "Regular Delivery Time." *See* UPS Initial Comments/Neels Report at 10 (Mar. 18, 2015) ("Initial Neels Report").

If the Postal Service's assumption that "Regular Delivery Time" is unaffected by the presence or absence of parcels on the route were correct, one would expect that, once parcel variables were included in the Regular Delivery Time model, the coefficients associated with those parcel variables would be zero: the number of parcels delivered should not have any effect on the regular delivery process. But that is not the

[3] Dr. Bradley confines his analysis to the question of whether the coefficients already in his regular delivery equation would change if parcels were included (*i.e.* "omitted variable bias"). This analysis does not examine the question of whether parcel volumes, if included in the model, would have a non-zero impact on regular delivery time. Dr. Neels analyzed the latter question and, as discussed herein, found including parcels does have a non-zero impact on regular delivery time.

case. Instead, the data shows a variability of 2.9% associated with parcel delivery, even on the segment of time the Postal Service treats as independent of parcel delivery and that excludes the "hands on" aspects of parcel delivery. *Id*.[4] Table 2 of Dr. Neels' initial report lists the full set of estimated coefficients, including the coefficients associated with parcels. *Id.* at 9.[5]

Finally, the Postal Service asserts that its own unreliable parcel data precludes it from including parcels in regular delivery time. *See* Postal Service Reply Comments at 20. But this only highlights the data limitations inherent in Proposal Thirteen, which attempts to develop cost attribution approaches for this large cost segment while using a small slice of data: 300 ZIP Codes out of a total of more than 10,000. These data limitations hardly justify compounding the problem by building unwarranted assumptions into the model. The solution to this problem, as Dr. Neels shows, is to use the much broader set of data available in the Form 3999 database. In addition, the Postal Service

[4] This result may be due, at least in part, to ambiguity about what "Regular Delivery Time" actually means. The Postal Service in its Reply Comments defines "Regular Delivery Time" as "the time associated with directly delivering letters and flats." Postal Service Reply Comments at 20. In the CCST Report, however, the Postal Service defined Regular Delivery Time more broadly as "includ[ing] primary delivery activities like driving along the route within delivery sections, accessing stops (whether on foot or in a vehicle), putting letters and flats into customers' mail receptacles, and retrieving collection mail from those receptacles." CCST Report at 19. This broader definition implicates all postal products including parcels. If that is so, then Regular Delivery Time *would* include activities beyond directly delivering letters and flats, which might explain why it is affected by parcel volumes.

[5] The Postal Service claims it is not assuming that "no street time is . . . associated with packages," but it is instead assuming the time associated with directly delivering letters and flats is not caused by those packages that are delivered separately from letters and flats. Reply Comments at 20. This argument attacks a straw man: UPS never claimed that "no street time, other than direct package delivery time" is associated with packages. UPS, instead, has shown that parcels affect so-called "regular" delivery time.

should fix the underlying issue by improving the data it collects going forward so that it properly accounts for parcels.

B. Proposal Thirteen Relies On Other Unwarranted Assumptions.

UPS demonstrated in its opening comments that Proposal Thirteen relies on the untested assumption that a certain (ambiguous[6]) cost component – Network Travel Time – is entirely fixed. Notably, the Postal Service's economist, Dr. Bradley, does not provide any statistical analysis justifying this assumption. Nor does the Postal Service. Instead, it asserts that, because it has long assumed that Network Travel Time costs are fixed, it should continue to do so. *See* Postal Service Reply Comments at 25-26. The Postal Service also tries to minimize this conceptual flaw by pointing out that Network Travel Time is not a large cost component.

Past classifications of what costs are fixed should not govern how the costs are treated today. This is especially true in light of the massive market-dominant volume declines in the last decade that have upended traditional thinking about fixed costs. *See* Charles McBride, *The Calculation of Postal Inframarginal Costs*, at 10 & Table 4 (2014) (expressing "serious reservations" about the Postal Service's methods of classifying fixed costs and noting that many of the Postal Service's "fixed" cost pools "declined even more than the system-wide total cost" during the massive volume declines since 2005).[7]

---

[6] Illustrating the complexities built into the Postal Service's approach, it remains unclear exactly what Network Travel Time represents. One would think Network Travel Time is the time needed to travel from one delivery point to another. But that does not seem to be the case. As noted, it appears that another activity – Regular Delivery Time – encompasses time traveling from one delivery point to another.

[7] In its forthcoming petition, UPS plans to address the Postal Service's treatment of "fixed" costs in a more systematic way extending beyond this segment.

In addition, as Dr. Neels notes in his Supplemental Report, Proposal Thirteen also assumes that the drivers of so-called "allied" activities are identical to the drivers of the purported "direct" street activities on which they have traditionally been piggybacked. *See* Supplemental Neels Report at 5-6. This assumption is yet another product of the Postal Service's fragmented approach to cost attribution.

In defending its approach, the Postal Service argues that what UPS has proposed is more fragmented than its own approach. *See* Postal Service Reply Comments at 27. In reality, Dr. Neels has developed a far less fragmented approach to cost attribution than Proposal Thirteen. In contrast to Proposal Thirteen, Dr. Neels' approach does not depend on carving city carrier costs into discrete activities and making assumptions about each. The National Form 3999 Model does not make assumptions about whether Network Travel Time is fixed or how allied and direct activities interact. It lets the data speak for itself.

C. Special Studies Are Burdensome And Prone To Error.

In its initial comments, UPS explained why the Commission should be encouraging costing approaches that do not rely heavily on special field studies. In responding to the comments of the Public Representative, the Postal Service acknowledges the tremendous burden special studies impose – even when those studies are limited because of budgetary reasons:

> A two week study involving 300 ZIP Codes dispersed throughout the country involved a tremendous amount of administrative resources. Not only is carrier time taken up in the data collection effort, but Headquarters, Area, and local personnel are necessary to assist with the study. Given the administrative resources available, collecting data from 300 sites over a two-week period approached the upper bound of cost. Trying to collect similar data for a year is not feasible. The cost of such a study should not be considered solely within the framework of the administrative cost, as

> the carriers who participated in the study have a job to do. Their job is to deliver the mail in a safe and timely manner. The carriers are a resource for the Postal Service, and their focus is to deliver mail. There is thus a limit on the duration of carrier involvement.

Postal Service Reply Comments at 8.

The expense and burden of the special studies is one reason why the Postal Service has managed to update its City Carrier dataset less than once per decade. An approach based on data collected in the ordinary course of business – such as that developed by Dr. Neels – could be updated on a more frequent basis. This is all the more important today given the rapid changes to the Postal Service's business.

Special field studies are also inherently artificial and prone to error. There is simply no way to be confident in the results of studies that ask mail carriers to try to time narrowly-sliced, discrete activities using the equivalent of stopwatches. While the Postal Service discusses some changes to how it conducted the studies between 2002 and 2014, improvements at the margins cannot cure this fundamental problem.

The special studies underlying Proposal Thirteen rely on self-reporting by the carriers themselves, which the Commission has long recognized lowers data accuracy. The Commission in 2005 identified carrier self-reporting, and the low-quality data it produces, as one of the main problems with the 2002 Study. *See, e.g.*, PRC Op. R2005-1, Appendix 1, at 11 ("The obvious downside to [carrier self-reporting] is that the data set may be quite inaccurate, since the quality controls exercised at the source are minimal by design."). As the Commission noted, "Proposed analyses do not improve upon established ones simply because the underlying data are more current. Data must not only be current, but be reasonably free of both sample error and reporting error." *See* PRC Op. R2005-1, at 54.

Reliance on self-reporting clearly affected the data quality in Proposal Thirteen. The CCST Report dedicates over twenty pages to how it excluded and scrubbed the carrier data. *See* CCST Report at 98-118. No matter how the Postal Service characterizes such efforts, it is clear that Proposal Thirteen requires extensive efforts to modify missing or erroneous self-reported data to produce workable results. *See, e.g.*, *id.* at 49 ("Because of the number of missing routes, these ZIP Codes were dropped from the analysis data set."); *id.* at 99 ("there were a substantial number of barcode pairs with zero recorded time," which forced Postal Service economists to make an "adjustment" to the data); *id.* at 94 ("a ZIP Code's data were included in the study data set only if the ZIP Code was able to provide at least one full week's worth of data").

The parcel study underlying Proposal Thirteen also had a high attrition rate. *See* CCST Report at 93-96 (indicating that about 4,500 routes participated in the final day of the parcel study, out of over 6,100 routes). The Postal Service tries to deflect concerns about this attrition in the parcel study by highlighting the higher participation rate of a *different study* – the *collection* study. *See* Postal Service Reply Comments at 33 ("UPS greatly exaggerates the level of 'study fatigue' that occurred and it ignores the fact the participation in the collection volume study was nearly universal."). That says nothing about the reliability of the parcel study, which is the key study under Proposal Thirteen for attributing costs to competitive products.

The Postal Service should replace its heavy reliance on costly and infrequent special field studies with using actual data collected in the ordinary course of business. In that regard, Proposal Thirteen is not a step in the right direction. Dr. Neels' Model, on

the other hand, shows how costing models for this segment can use actual operational data. This is a major difference between the two approaches.

D. Special Purpose Routes

In its initial comments, UPS pointed out that Proposal Thirteen does not address the treatment of Special Purpose Routes and explained why these routes deserve closer scrutiny by the Commission. This is particularly true given that UPS understands the Postal Service is increasingly using special purpose routes to deliver parcels, driven by growth in the competitive side of the business.

In response, the Postal Service notes that 56% of the *attributable* street costs of Special Purpose Routes are attributed to competitive products. *See* Postal Service Reply Comments at 39. But this glosses over the fact that only 42% of total costs in this component are attributed, meaning (apparently) that just *23%* of Special Purpose Route total costs are attributed to competitive products. In light of the increasing use of Special Purpose Routes to deliver competitive products, there are serious questions about whether these numbers make sense.[8]

The problem here is not just one of accuracy; it is also one of transparency. Despite spending a great deal of effort combing through costing files, and despite raising these issues in its initial comments, UPS still does not have clarity from the Postal Service about how it is accounting for the costs of Special Purpose Routes. It

[8] The treatment of Special Purpose Routes suffers from conceptual flaws stemming from the Postal Service's narrow definition of cost causation. Because over 90% of parcels are competitive products, the overwhelming majority of costs associated with routes that exclusively or almost exclusively deliver competitive products should be attributed to competitive products. To label a cost as institutional when over 90% of the associated products are competitive creates the perverse result that competitive products must only bear 5.5% of these costs, through the "appropriate share" requirement.

remains unclear, for example, what fraction of the costs of Special Purpose Routes are associated with routes that deliver only (or almost only) competitive products. This prevents the Commission from assessing the reasonableness of the variability assumptions the Postal Service is making.

The Postal Service should, at a minimum, be required to provide clarity about these issues. Specifically, the Postal Service should be directed to answer the following questions:

(1) What portion of the costs included in the Special Purpose Route cost component is associated with parcel-only routes?

(2) How has this fraction changed over the past five years?

(3) How has the number of individual parcel-only routes changed over the past five years?

(4) What specific studies or assumptions provide the basis for determining the division of the Special Purpose Routes cost component between institutional costs and attributable costs?

The answers to these questions will help the Commission and others assess whether Special Purpose Routes are being handled in an appropriate manner.

III. THE COMMISSION SHOULD ADOPT DR. NEELS' NATIONAL FORM 3999 MODEL.

The Commission should not adopt Proposal Thirteen in light of its flaws. The use of special studies can and should be retired in favor of using national Form 3999 data, which is collected in the ordinary course of business. Artificial divisions of city carrier time into buckets for parcel and "regular" delivery time can and should be abandoned in favor of a holistic approach. The Postal Service should not slice its costs into a series

of artificial and often ambiguous components and then make assumptions about how they behave. The data should be allowed to speak for itself. Dr. Neels has developed a model – the National Form 3999 Model – that achieves these goals.

The framework of the National Form 3999 Model draws from the best features of Proposal Thirteen. As discussed in his Supplemental Report, Dr. Neels utilizes a similar model specification to the one proposed by the Postal Service in Proposal Thirteen: a flexible form model that contains linear and squared terms for each postal product as well as cross-terms to model the interactions between and among postal products.

In his initial report, Dr. Neels criticized Proposal Thirteen's use of this type of specification as too complex, pointing to its numerous insignificant coefficients and imprecisely-estimated marginal costs. Dr. Neels observed that these manipulations raise a concern that the special study datasets would not support a complex specification containing such a multitude of explanatory variables. That criticism arose from the fact that the underlying data supporting the Postal Service's model is too limited in scope and extent.

With the additional data made available using the Crosswalk File, however, Dr. Neels can use the full set of Form 3999 data. Since the Form 3999 dataset is significantly broader in scope, it can support a more complex model specification. While Proposal Thirteen's modeling framework was too complex for the limited Proposal Thirteen dataset, it is not too complex for the broader Form 3999 dataset.

After receiving the Crosswalk File, Dr. Neels conducted a series of tests that demonstrated that the flexible functional form is a better fit once the broader set of data is used. As Dr. Neels explains:

> [Dr. Neels' calculations] provide compelling evidence that the flexible quadratic form provides a significantly better description of the relationship between gross street time and the available measures of mail volume and the mail delivery environment than the nonlinear model presented in my initial report. The substantially larger estimation dataset provided by the Form 3999 data permits me to estimate the many coefficients of this model with a high degree of precision. In order, therefore, to provide the most accurate possible estimates of the relationships between mail volume and street time, and to eliminate a source of disagreement and controversy between myself and the Postal Service, I will use the flexible quadratic form for the remainder of my analysis.

Supplemental Neels Report at 26-27.

Thus, in his Supplemental Report, Dr. Neels uses the modeling specification favored by the Postal Service for this cost segment (the flexible functional form) but applies it to the full set of Form 3999 data, instead of the limited set of 300 ZIP Codes used in Proposal Thirteen.

Dr. Neels also removes the unwarranted assumptions the Postal Service had grafted onto that specification. To do this, Dr. Neels incorporates parcel-related variables into the modeling framework, removing the artificial distinction between parcel and regular delivery time embodied in Proposal Thirteen. Thus, Dr. Neels' model accounts for possible interactions between parcels and any of the other postal products.

For similar reasons, Dr. Neels models city carrier street time as a whole.[9] Instead of carving city carrier street time activities into a number of discrete activities and positing a different model for each activity, Dr. Neels models the entirety of a city carrier's daily route as one bucket. This approach evaluates variability of the cost segment all at once, instead of artificially dividing street time into many buckets that are

[9] The exceptions are that accountable time and collection box time have been subtracted from gross street time. Dr. Neels explains the rationale behind the decision to exclude these two small portions of time in his Supplemental Report. Supplemental Neels Report at 27-28.

handled independently of each other. It does not require making assumptions about whether certain activities are fixed or how one activity affects another. It dispenses with having to assume that the drivers of allied activities are identical to the drivers of the direct street activities. To the extent this assumption is appropriate, Dr. Neels' model confirms it, and to the extent it needs to be adjusted, the model adjusts it appropriately.

Together, these changes yield a model that is conceptually superior to Proposal Thirteen. The model generates strong statistical and econometric results, and it attributes a more realistic share of costs to competitive products. Accordingly, UPS submits that the Commission should adopt Dr. Neels' National Form 3999 Model.

IV. THE POSTAL SERVICE SHOULD IMPROVE ITS COLLECTION OF FORM 3999 DATA.

The Commission should also direct the Postal Service to collect as part of its regular route evaluation process reliable volume data for all of the mail streams for which costs need to be attributed. This would mean augmenting the Form 3999 dataset to include accurate volume measures for accountables, collection volume, deviation parcels, and in-receptacle parcels. The highest priority should be placed on improvement of the competitive volume data collected by the Postal Service.

This request should not be overly burdensome for the Postal Service. The route evaluations that make up the 3999 dataset occur for a typical route approximately once every three or four years. It is not too much to ask that, when such an evaluation occurs, it should include accurate mail counts for all mail streams, including the parcel stream that accounts for an increasing share of the Postal Service's business. Given the frequent statements by Postal Service management regarding the extent to which

parcels will be the mainstay of its future business, the importance of improving the quality of parcel count data is self-evident.

The availability of improved volume count information will improve the accuracy of the costing parameters derived from the Form 3999 dataset, and support moving away from special studies, establishing a system that can be updated annually and kept abreast of rapidly changing trends in the delivery environment. Similarly, more accurate Form 3999 data will permit a better understanding of the costs associated with mail collection and accountables.

As Dr. Neels explains, because of the Postal Service's inadequate collection of parcel-related data as well as data on customer collection volume, the National Form 3999 Model must for now rely upon imputed volume measures for a few mail streams. Dr. Neels has developed imputation approaches that he discusses at length in his Supplemental Report. *See* Supplemental Neels Report at 27-37. Although these approaches are not perfect, they are sufficient for now and can be replaced with actual data in the future once the Postal Service begins collecting better data. This adaptability is yet another strength of the National Form 3999 Model.

If the Commission is unwilling to accept the National Form 3999 Model now, either because of these data limitations or any other reason, the Commission should require the Postal Service to modify Proposal Thirteen to include parcels in the equation for "regular" delivery time. Dr. Neels demonstrated in his initial report that including parcels in the regular delivery time specification leads to the conclusion that there is a 2.9% parcel variability associated with that cost pool, which the Postal Service incorrectly assumes to be independent of parcels altogether.

The fact that Dr. Neels obtains non-zero parcel coefficients when analyzing a bucket of time that is supposedly completely independent of parcel delivery time demonstrates that those buckets are not actually independent. Even if the Commission does not adopt UPS's alternate approach, it should at a minimum ensure the Postal Service modifies its city carrier model to remove an assumption the data does not support. Relaxing this incorrect assumption increases overall variability of "regular delivery time" from 35.5% to 36.2%. As Dr. Neels explains, this would be a conservative modification. *See* Supplemental Neels Report at 43.

V. CONCLUSION

UPS respectfully requests that the Commission adopt the National Form 3999 Model developed by Dr. Neels for this cost segment. This will improve accuracy and transparency, and it will better meet the congressional goals set forth in PAEA. The Commission should also require the Postal Service to answer the questions identified above regarding Special Purpose Routes. Finally, the Commission should require the Postal Service to improve data collection methods regarding competitive products.

Respectfully submitted,

UNITED PARCEL SERVICE, INC.,

By: /s/ Steig D. Olson

Steig D. Olson
Quinn Emanuel Urquhart & Sullivan, LLP
51 Madison Ave., 22nd Floor
New York, NY 10010
(212) 849-7152
steigolson@quinnemanuel.com
Attorney for UPS

Exhibit A

BEFORE THE

POSTAL REGULATORY COMMISSION

WASHINGTON, DC 20268-0001

PERIODIC REPORTING
(PROPOSAL 13)

Docket No. RM2015-7

SUPPLEMENTAL REPORT OF
KEVIN NEELS
ON BEHALF OF
UNITED PARCEL SERVICE

Table of Contents

I. Introduction

This report is the second report I am filing on behalf of UPS in Docket RM 2015-7, in which the Postal Service has proposed updates to its costing model for city carrier street time. In my March 18, 2015 report, ("initial report") I outlined a number of flaws in the Postal Service's econometric study of city carrier street time variability.[1] I also urged the Commission to consider an alternative approach that addresses many of these flaws and that would allow for more regular updates to the city carrier street time model by leveraging data regularly collected by the Postal Service and by decreasing reliance on costly special studies. My initial report also contained an illustrative model following such an approach that, while necessarily incomplete, demonstrated promise and resulted in variability estimates, and thus had costing implications, that in many ways differed significantly from those in the Postal Service model.

On April 23, 2015, the Commission issued an order granting access to a confidential zip code crosswalk that would allow for an exploration of the suitability of the national Form 3999 dataset for such an alternative approach.[2] The national Form 3999 dataset contains most of the variables used by the Postal Service in its model, but is substantially larger, providing these variables for every city carrier route nationwide. In that same order the Commission established a revised schedule that provided UPS with an opportunity to file this supplemental report based on my analysis of the new materials described above.

On May 13, 2015, the Postal Service and a number of other parties filed comments, some of which were directed at my initial report. Among the materials filed by the Postal Service was a report by Professor Michael Bradley of George Washington University that contained extensive comments on my initial report. [3]

In this Supplemental Report, I present further findings on the feasibility of using the national Form 3999 data to pursue such alternative approaches for the analysis of city carrier costs. I present models that more closely align with the approach espoused by the Postal Service, yet have costing implications that depart in significant ways from the model proposed by the Postal Service. I also respond to criticisms made by Professor Bradley and the Postal Service of my initial report. Finally, I

[1] Report of Kevin Neels on Behalf of United Parcel Service, Docket RM2015-7, March 18, 2015 ("Initial Neels Report"). Report on the City Carrier Street Time Study, Docket RM2015-7, December 11, 2014 ("Street Time Study Report").

[2] Order No. 2455 - Order Granting United Parcel Service, Inc. Motion for Issuance of Commission Information Request No. 1 and Revising Procedural Schedule, Docket RM2015-7, April 23, 2015.

[3] Analysis of the Report of Dr. Kevin Neels On Behalf of United Parcel Service, Docket RM2015-7, May 13, 2015 ("Bradley Reply Report").

provide the Commission with recommendations to consider both with respect to the city carrier costing model that should be used in ACR2015 as well as how to facilitate the continuing improvement of the city carrier costing model going forward.

One of my goals in undertaking this research was to find a way to reduce the need for the Postal Service and the Commission to rely on costly and infrequently updated special studies. Original data collection is expensive, and for this reason the size, scope and frequency of special studies will always be heavily influenced by budget constraints.[4] These predictable constraints are greatly exacerbated by the Postal Service's ongoing financial difficulties. A strategy that is based upon the results of costly special studies runs a chronic risk of being forced to rely on analyses that are under-powered, over-simplified and out of date.

A second important goal was to develop an approach that would account properly for the growing importance of parcels. While the Postal Service has experienced historical declines in overall mail volume in recent years, the number of parcels that it delivers has grown sharply. In numerous statements the Postal Service has made it clear that it believes that in the future, delivery of parcels will be one of the primary if not *the* primary focus of its business. Taken together, these facts imply that the role of parcels is changing substantially and rapidly. One of my goals is to try to make sure that these changes are captured accurately by Postal Service costing procedures.

My third goal was to devise an approach that would examine city carrier delivery activities as a whole, rather than as a collection of distinct and separate activities that are examined, measured and analyzed in isolation. As I explained in my initial report, I am not confident of the Postal Service's ability to measure accurately the amounts of time that city carriers spend carrying out these various discrete activities. The 'product' of delivering different types of mail is jointly produced by a mail carrier. Consequently, I also suspect that these various activities are sufficiently intertwined as to make it impossible, when they are viewed in isolation, to understand and measure accurately the factors that drive their costs.

These three goals are clearly interconnected. The rapid changes taking place in the size and composition of the mail stream increase the importance of moving away from costly and infrequently updated special studies and toward a new approach promising fresher and more relevant results, and an improved ability to track in a timely manner the rapid changes taking place in the delivery environment. My desire to adopt a more holistic approach to the analysis of city carrier delivery costs is driven to a large extent by my skepticism regarding traditional Postal Service assumptions

[4] In its Street Time Study Report the Postal Service freely admitted that its decision to collect data for a sample of 300 zip codes was dictated by "budgetary and management resources," rather than by what was required to achieve threshold levels of statistical precision or meet other experimental design goals. See Street Time Study Report, p.28.

that regular delivery activities are entirely unaffected by the presence of parcels, and that any additional costs associated with the delivery of parcels are incurred only when a letter carrier is standing in front of a mail box with a parcel in hand.[5] Adoption of a more holistic modeling strategy that takes overall street time as the dependent variable, and bases inferences about cost causation on statistical relationships offers a valuable opportunity to test the validity of these assumptions.

The remainder of the report will begin with a summary of my findings. I then address some of the comments and criticisms directed at my initial report by the Postal Service and Professor Bradley. In addressing these comments I attempt to identify and clarify for the Commission what I think are the principal points of disagreement between us. This section is followed by a description of my analysis of city carrier cost variability using the national Form 3999 dataset. I briefly discuss the central results of this analysis, and then conclude by presenting a set of recommendations for the treatment of city carrier costs both in the near term and the longer term.

II. Summary of Findings

I have evaluated the national Form 3999 dataset, and concluded that it can be used to provide accurate and reliable measurements the effects of changing mail volumes on city delivery costs in a form suitable for ongoing use in product level cost determination.

I have investigated the deficiencies in the national Form 3999 dataset – specifically, the absence from this dataset of accurate volume measures for a number of important mail streams, and have concluded that these deficiencies can be rectified. I have identified methods for estimating the missing volume measures that appear to give reliable results and yield volume estimates that appear to perform well in statistical analyses.

Based upon a careful evaluation of the alternatives, I have combined what I believe to be the best features of the model presented by the Postal Service in its Street Time Study Report, and in my March 2015 report. This model focuses on overall street time, includes as explanatory variables volume measures for all of the major mail streams as well as delivery environment variables drawn from both models, and employs a flexible functional form. Estimation of this model using the full

[5] In its Reply Comments the Postal Service argues that it does not assume that this general street time is not associated with packages, but rather is assuming that the time associated with directly delivering letters and flats is not caused by those packages that are delivered separately from letters and flats. See Reply Comments of the United States States Postal Service In Response to March 18th Comments, Docket RM 2015-7, May 13, 2015 ("Reply Comments"), p. 20. However, the measure of time spent "directly delivering letters and flats" used in the Postal Service's cost analysis is Sector Segment Time. See Street Time Study Report, p. 9. Sector Segment Time includes the time spent traversing segments of the carrier's route, See "Performance And Field Operations Support Data Collection Device Street Activity Standardized Definitions," pp. 4-7.

nationwide Form 3999 dataset produces precise measurements of volume variability for all of the major mail streams.

I discuss four potential choices regarding the city carrier costing procedures to be employed in the Annual Compliance determination for Fiscal Year 2015. These include maintenance of current procedures, adoption of the Postal Service's Proposal Thirteen, adoption of a modification of Proposal Thirteen based upon econometric results presented in my March 2015 report that assigns a portion of regular delivery time to deviation parcels, and a costing procedure based upon the econometric analysis based on the nationwide Form 3999 dataset described above. I present calculations showing the attribution of city carrier costs to products that would result from adoption of each of these four alternatives.

I recommend that the Commission adopt the costing procedures based upon my analysis of the nationwide Form 3999 dataset for use in the Annual Compliance Determination for Fiscal Year 2015, recognizing that this best alternative, like all of the others, is currently limited by the quality of the data collected by the Postal Service. If the Commission decides against adoption of this alternative, I recommend adoption of my proposed modification of Proposal Thirteen.

I recommend that in order to provide a sound empirical basis for future cost studies in this area, and to eliminate the need for future costly special studies the Commission should direct the Postal Service to improve the quality of the parcel volume data collected during the route evaluations whose results are reported in the Form 3999 dataset, and to explore the feasibility of measuring collection mail volumes.

III. Postal Service Comments on My Initial Report

Given the extensive comments on my initial report filed by the Postal Service and by Professor Bradley, I think it is appropriate to begin by responding to some of the points raised therein, as they bear directly on the rationale for a number of the analytical decisions I made in the course of conducting the research reported below.

Professor Bradley poses a number of questions in his report. Having considered these carefully, I think that they relate to three fundamental issues on which I and UPS on the one hand, and Professor Bradley and the Postal Service on the other, have genuinely divergent views. One has to do with the extent to which parcels can be segregated from regular delivery and analyzed separately. A second has to do with the extent to which variations in mail volume over the days of the week can provide insights into cost causation that should guide the deliberations of the Commission. The third has to

do with the adequacy of the sample sizes provided by the special studies conducted by the Postal Service. [6]

Let me immediately make it clear that in questioning the adequacy of the special study samples I in no way meant to imply that the work carried out by the Postal Service in connection with these studies was careless or ill intentioned. Clearly the Postal Service devoted substantial time, energy and resources to these efforts and their results represent an improvement over the prior studies that formed the basis for the currently accepted city carrier costing parameters and methodology. Nonetheless, precisely because of the resources that these studies required, there was a limit to the volume of data that they produced. For this reason the questions of just how much information they produced and how complex a modeling effort that information will support have to be considered seriously.

A. Including Parcel Volumes in the Regular Delivery Equation

A central theme of my initial report was that the separation of parcel and accountable time from "regular" delivery time was artificial, and that the assumption that parcel volumes have no effect on regular delivery times needed to be tested.[7] If it is the case, despite implicit assumptions to the contrary made by the Postal Service, that parcel volumes have a demonstrable effect on regular delivery time, this would mean that Proposal 13 does not attribute sufficient costs to parcels and the associated postal products.

Professor Bradley has taken issue with certain tests of the December 2014 Postal Service regular delivery model that I performed in order to analyze this very question. [8] In these tests I included parcel volume variables in the regular delivery time model. Professor Bradley's comments on these tests are framed very narrowly in terms of whether or not the coefficients of the non-parcel terms in the regular delivery equation suffer from omitted variable bias. He concludes that the statistical evidence is insufficient to prove the existence of such bias.

I think this issue and the nature of my disagreement with Professor Bradley runs deeper. I understand that a certain amount of indirect time is piggybacked onto other directly costed activities.

---

[6] Professor Bradley also spends considerable time and effort in his reply report critiquing and extending a result contained in my initial report that challenged the statistical precision of the regular delivery model used in Proposal 13. He uses these calculations to criticize my alleged support for the aggregation of volume measures within the context of the regular delivery time model used in Proposal 13 – a position that I do not actually espouse. See, for example, Bradley Reply Report at pp. 14-27 . Another set of Professor Bradley's critiques (which I disagree with) are irrelevant, given the direction my analysis has taken. See, for example, Bradley Reply Report at pp. 34-39. Without conceding the validity of these criticisms, in the interest of the Commission's time, I will not take time here address them.

[7] See Initial Neels Report, pp. 5-11.

[8] See Bradley Reply Report, pp. 10-14.

However, within the directly attributed city carrier costs, the only costs attributed to parcels comes from the "parcel time" taken from the national Form 3999 dataset and adjusted based on the parcel delivery time data collected in the special parcel study. As I understand the special parcel study documentation and training materials, it appears that most of the parcel time captured in that study was incurred near the mail receptacle or, in the case of deviation parcels, going to and from the recipient's door. If that understanding is correct (and I welcome clarification on this point), then the analysis presented in the Street Time Study Report assumes that the presence of parcels has no direct effect on any other aspect of the delivery process. This may be a strong assumption, especially given the growth that has taken place in parcel volumes, and that is likely to continue to take place. At a minimum, I believe, the validity of this assumption should be tested.

More generally, however, I understand that over the course of his or her day a letter carrier handles a number of different mail streams. Parcels make up one of those streams. Delivering different streams of mail is a jointly produced service and the handling of these streams is at various points of time thoroughly intermingled. This fact causes me to question whether it is even possible to isolate and measure the time spent dealing with a single mail stream. There may certainly be instances over the course of the day when the letter carrier is handling only one type of mail. However, I would expect those times to add up to much less than the full day. During some major portion of a carrier's day he or she will be dealing with multiple mail streams. It is reasonable to assume that the time required to carry out these other activities is potentially influenced by all mail volumes, to an extent that can probably only be measured statistically. In my opinion, it is not possible to manually unscramble the eggs in a situation like this in order to figure out how much time you spent cooking the egg whites.

In my view, there is an obvious way to address these concerns. This can be done by taking total delivery time as the dependent variable in the model, and allowing the results of the statistical analysis to measure the marginal costs associated with each individual mail stream. These methods are well suited to sorting out complex patterns of cost causation, such as those like the city carrier delivery environment. The results presented below are based on that approach.

The results of the regression provided in Table 2 of my initial report provide a more direct test of the assumption that parcels do not affect regular delivery time as that term is defined by the Postal Service in Proposal 13. The fact that several of the parcel coefficients are positive and that the point estimate of marginal costs are economically significant suggests that parcels do affect regular delivery times and undermine the validity of that assumption.[9]

[9] Using heteroscedasticity robust standard errors one can convincingly reject the null hypothesis that the coefficients in this equation involving parcel volumes are jointly equal to zero. Using clustered standard errors the calculated probability approaches but does not quite reach the 10 percent critical value for rejection of the null hypothesis. Given the small sample size upon which the test was based, the known problems with the

On a narrow technical note, I respectfully disagree with Professor Bradley's statement that the statistical comparison presented in Table 2 on page 13 of his report demonstrates that there is no omitted variable bias in the regular delivery model. He tests the null hypothesis of equality of the estimated non-parcel coefficients between the models with and without parcel variables included, and fails to reject that hypothesis. Failure to reject the null hypothesis of equality is not the same thing as proving equality. Instead, all that he has shown is that we can't tell. In my view, this inconclusiveness is another manifestation of the fact that the special study sample size is too small to support firm conclusions, a point I discuss in more detail below. Furthermore, the failure of the inclusion of parcels to generate statistically significant changes in the variabilities and marginal costs of DPS, SEQ, FSS, cased mail, and collection volumes is at best ancillary to the question at hand. That question – whether the volume of parcels affects regular delivery time and is a relevant driver of an economically meaningful amount of costs that would otherwise be treated as institutional or be attributed to other mail streams – is best answered by the tests I have proposed and performed. They show that parcel volumes do affect regular delivery times.

B. Can Day of the Week Volume Changes Be Used to Measure Volume Variabilities?

In my initial report I noted that much of the variation in volume contained in the special study dataset consisted of small variations in volume over the successive days of the week, and I criticized the study for having an excessive short-term focus.[10] Professor Bradley appears to have been confused by my use of the phase "longer term" in characterizing what I believe to be the appropriate variabilities to estimate and employ for costing purposes.[11] I take this opportunity to clarify that I use the terms "longer term" and "longer run" interchangeably, and to restate the importance and relevance of this issue to the decisions facing the Commission in this docket.

Professor Bradley correctly characterizes the "short run" as a situation in which one or more inputs to the production process is constrained from adjusting to changes in prices or output levels. He argues, again correctly, that one would in general expect short run cost variabilities to exceed long run cost variabilities. For example, a firm faced with a sudden surge in demand for its product might decide to accommodate this new demand by extending the work day of its employees, incurring in the process

---

parcel volume measure used in the test (i.e., DOIS parcel volumes), the suggestive results of the test despite these shortcomings and even when it is based upon clustered standard errors, and my strong prior belief that parcels should have an effect on regular delivery time, I accept that the parcel variables belong in the equation, and that the estimated coefficients represent (as statistical principles indicate) the best available estimate of their effects.

[10] See Initial Neels Report, p. 14.

[11] See Bradley Reply Report, p. 5. I am using the term "variability" in the same sense that it is used by the Commission and the Postal Service. In this context, it is defined as the elasticity of cost with respect to a cost driver.

high overtime costs. If the surge in demand proved permanent, given more time, that same firm might decide to hire additional workers, allowing it to achieve the same level of production at a lower overall cost. Finally, Professor Bradley notes, again correctly, an exception to this general expectation occurs when there is excess capacity in the system.[12]

Having established these areas of general agreement, I can now clarify the points I sought to make in my initial report, explaining in the context of city carrier delivery what the relevant differences are between long run and short run variabilities, and the nature of the excess capacity that exists within the system.

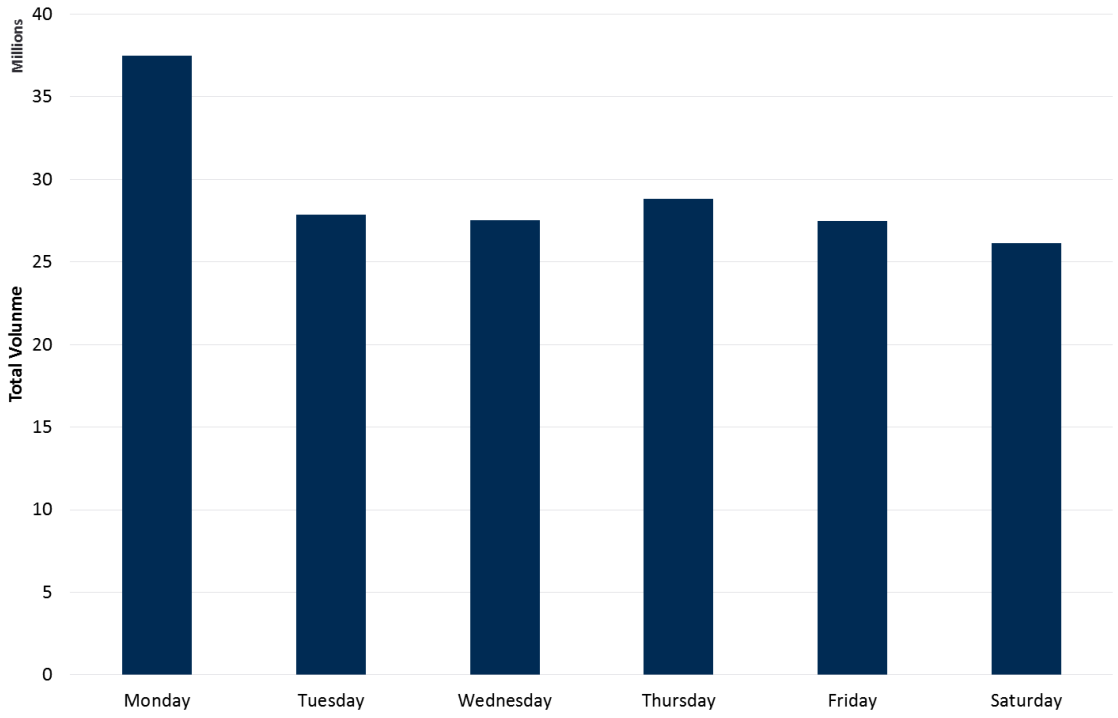
In the context of city carrier delivery the most important distinction between long run and short run variabilities relates to the presence or absence of route restructuring. In numerous filings, reports, and pieces of testimony, the Postal Service has explained that it routinely responds to changes in mail volume by restructuring delivery routes in an effort to save costs and keep route-specific workloads to a level in which they fit within a normal workday. For example, the introduction to the Street Time Study Report states that one of the major changes that has occurred in the delivery environment is "the restructuring of the city carrier network."[13] In the context of city carrier delivery it is appropriate to define the short run volume variability of delivery time as the change in delivery time associated with changes in volume, holding route structure fixed, and the long run variability as the changes in street time associated with changes in volume, allowing route structure to adjust.

Equipped with these definitions, it is also easy to identify the excess capacity that exists in the system, and that allows long run volume variabilities to exceed short run volume variabilities. That excess capacity is the result of predictable day to day variation in mail volume. Figure 1 summarizes variation in mail volumes by day of the week within the dataset collected by the Postal Service for the City Carrier costing study. The pattern of variation in volume is clear. Volumes are highest on Mondays and consistently lower for the remaining days of the week. One has to assume that the Postal Service is aware of this variation, and takes it into account when it restructures delivery routes.

[12] See Bradley Reply Report, p. 9. Professor Bradley asserts that if there is excess capacity in the system, marginal costs and variabilities would be equal to zero. I don't believe that this argument is correct. There could be excess capacity in the sense that the volume of one or more inputs to the production process is fixed at higher than optimal levels, creating a situation in which the quantities of non-fixed inputs can be reduced, but not necessarily to zero.

[13] See Street Time Study Report, p. 1.

Figure 1: Delivery Volumes by Day of the Week
As Shown in the Special Parcel Study Data



Notes and Source: Sum of DOIS mail volumes by day of week for all route day records used to estimate the USPS RM2015-7 Model.

The fact that mail volumes routinely vary over the course of the week means that in restructuring city carrier delivery routes the Postal Service faces a well-known problem that frequently arises in transportation networks. When volume varies, how much capacity should one provide? Answering this question involves making tradeoffs between the amount of congestion and delay one is willing to accommodate during peak periods, and the amount of additional cost associated with capacity that goes unused during slack periods that one is willing to bear. The correct answer almost always involves providing a level of capacity that falls somewhere between the highest and lowest volumes one expects to encounter. The result, inevitably, is that during slack periods there is some degree of excess capacity.

In response to a sustained drop in volume the Postal Service is likely to restructure its network of routes in order to achieve more efficient utilization of delivery personnel and to reduce route-related overhead costs. I would not expect it to carry out a similar restructuring over the days of the week, even though there are substantial volume differences across these time periods.[14] Hence, on low

[14] I recognize that the Postal Service may make some day-to-day adjustments to delivery procedures, but would not expect them to be as extensive as those that emerge from the periodic route restructuring process.

volume days of the week, there is likely to be excess capacity. Relative to the Monday peak, regular delivery times will be lower on Friday, but not to the same extent that one might expect to see in response to a sustained and across the board drop in volume of comparable magnitude.

The ideal data one would need to measure long run variability as I have defined it would be zip-code level delivery costs measured over the course of several route evaluation cycles taken over a course of several years. Operationally, as I understand Postal Service data collection, such a dataset would consist of several snapshots of the Form 3999 data over the space of several years and pooled for analysis purposes. In the absence of such data, the nationwide cross-section of data as collected in the Form 3999 data, which captures a wide range of delivery environments and volume intensities, appears to be the best available proxy. Estimating a cost variability model from panel data covering a relatively small number of zip codes, in which the cross-sectional variation is accompanied by variation over the successive days of a two week period, results in lower estimates of volume variability relative to the cross-sectional model, as evidenced in Tables 6 and 7 of my initial report.[15]

Relying for variability measurement purposes on a data set in which much of the volume variation consists of short run day of the week fluctuations in mail counts produces, in my view, an estimate of the wrong measure – one that is too heavily tilted toward short run responses of delivery time to small changes in volume in the context of a route structure that is essentially fixed. I believe that for purposes of establishing cost causation the Commission should focus on the long-run response of street time to changes in volume.

Professor Bradley's report contains an extensive discussion of the lack of precision and statistical significance of the coefficients of the version of the Postal Service model that I estimated using zip code level averages of the volumes collected in the special study. I don't deny that the cross-sectional model presented in my original report suffers from these defects. Indeed, throughout that report I repeatedly expressed concern over the extent to which the ambitious specification upon which the Postal Service model was based taxed the limits of the small volume of data collected in the special study. In the case of the cross-sectional model, which sought to estimate over 25 separate parameter values from a dataset containing less than 300 observations, such concerns were greatly amplified.

My point in presenting the cross-sectional model was certainly not to suggest that it provided robust results that the PRC and the Postal Service could rely upon. Rather, my goal was to show that if one ignores the day of the week related volume changes that make up so much of the total volume variation in the special study dataset, one tends to arrive at higher volume variability estimates.

[15] See Initial Neels Report, pp. 10-14.

Professor Bradley challenges this conclusion, arguing that if one eliminates the many variables that take insignificant coefficients in the cross-sectional model, the resulting model implies lower rather than higher volume variabilities. Respectfully, I disagree with his conclusions on this point. Eliminating all of the variables whose coefficients fail to achieve conventional significance levels doesn't leave one with much of a model.[16] The resulting model is likely badly misspecified, and incapable of producing reliable results.

C. Do the Special Studies Provide an Adequately-Sized Sample?

Finally, I consider the question of the adequacy of the sample sizes provided by the special studies conducted by the Postal Service. In one sense they appear to be fairly large, containing thousands of observations. In an important sense, however, they are relatively small, providing information for a sample of at most 300 zip codes. The larger sample size is generated by taking repeated observations of those same 300 zip codes on successive days over a two week period. I differ from the Postal Service and from Professor Bradley in my assessment of the amount of useful information provided by those repeated observations. In the section above I argued that their contribution was of limited relevance because they skewed the results of the study toward an inappropriately short run measure of variability.

However, quite apart from the question of whether the time series dimension of the special studies dataset is *conceptually* relevant, there are also reasons to question the extent to which it is *statistically* relevant. These reasons can be illustrated by a thought experiment. Suppose that it were the case that mail volumes and the other delivery environment variables did not change at all from one day to the next. In this Groundhog Day-like world we would repeatedly observe the response of delivery time to exactly the same set of stimuli. Increasing the sample by adding additional days would provide us with an increasingly accurate measurement of the amount of natural variation that occurs in delivery time. But we would never have the opportunity to observe how delivery time responds to new volume levels or delivery environments.

Now take the thought experiment one step further, and suppose it is also the case that the random factors that influence delivery time are correlated over time. Such a situation could easily arise from idiosyncratic aspects of a location that are not captured by the explanatory variables of the model that cause the model to consistently under or over-predict delivery time. In such a case, adding additional days of data to the sample would not necessarily even improve our estimates of the amount of natural variation in delivery time that occurs.

---

[16] I further note that 8 of 22 coefficients in the final IR Package Delivery Time Model (see Street Time Study Report, p. 112) and 17 of 30 coefficients in the final Deviation Delivery Time Model (p. 116) did not achieve conventional significance thresholds, a fact which did not prevent the Postal Service from presenting them to the Commission as key elements of Proposal 13.

Situations such as that which I have described arise frequently in empirical research, and there are well accepted techniques for dealing with them. Just as there are ways of calculating standard errors that account for the presence of heteroscedasticity, there are also methods for calculating standard errors that account for other departures from the standard assumptions upon which ordinary least squares are based. In particular, when (as is the case for the special study dataset) observations fall naturally into groups within which error terms are likely to be correlated, there are methods of accounting for this clustering effect that provide unbiased standard error estimates.[17]

I calculated clustered standard errors accounting for within zip code correlations for the regular delivery model presented in the Street Time Study Report. The t-statistics resulting from this calculation are presented in Table 1, which also contains the t-statistics presented in the original report. The final version of that model included 27 estimated coefficients, all of which were judged in the Report to be statistically significant. However, when one evaluates these coefficients using clustered standard errors, a quite different picture emerges. Accounting appropriately for the clustering of errors within zip codes causes 8 of the 27 coefficients in the model to fail to achieve statistical significance. These imprecisely estimated coefficients include the FSS dummy, the first order term for cased mail, the first and second order terms for collection mail, the interaction terms between cased mail and delivery points and collection mail and delivery points and both of the business ratio variables.

[17] Some econometricians believe that in practice clustering is a much more serious problem than failure to account for heteroscedasticity, a subject that is extensively discussed both in the Street Time Study Report (pp. 56-65) and in the Bradley Reply Report (pp. 36-38). For example, Angrist and Pischke write that "Bias problems aside, heteroscedasticity rarely leads to dramatic changes in inference. In large samples where bias is not likely to be a problem, we might see standard errors increase by about 25 percent when moving from the conventional to the HC1 estimator. In contrast, clustering can make all the difference," Joshua D. Angrist and Jörn-Steffen Pischke, "Mostly Harmless Econometrics: An Empiricist's Companion," March 2008.

Table 1: The Postal Service's Proposal 13 Regular Delivery Regression Results with Clustered Standard Error T-Statistics

Variable	Estimated Coefficient	Proposal 13 Reported T-statistic	Clustered T-statistic
INTERCEPT	(18.2)	(12.42)	(4.23)
FSS Dummy	**3.86**	**2.86**	**1.27**
DPS	1.81	4.04	2.49
DPS2	(0.0000244)	(5.33)	(3.44)
CM	**2.96**	**2.38**	**1.31**
CM2	(0.0000746)	(3.10)	(2.29)
SEQ	3.33	8.77	6.02
SEQ2	(0.0000736)	(5.69)	(3.68)
FSS	8.40	5.98	2.92
CV	**4.06**	**2.14**	**0.77**
CV2	**(0.000295)**	**(3.25)**	**(1.18)**
DP	24.5	23.69	9.14
DP2	(0.000472)	(8.42)	(3.70)
DPS*CM	0.0000683	3.03	1.96
DPS*CV	(0.000223)	(4.89)	(3.04)
DPS*DP	0.000154	4.31	2.14
CM*CV	0.000384	3.99	2.14
CM*DP	**(0.000180)**	**(2.38)**	**(1.20)**
FSS*CV	0.000445	4.29	2.14
FSS*DP	(0.000387)	(5.05)	(2.31)
CV*DP	**0.000489**	**3.86**	**1.42**
DM	45.5	14.63	4.78
DM2	(27.4)	(8.56)	(2.80)
MPDP	79.4	6.51	2.12
MPDP2	(136)	(6.57)	(2.13)
BR	**(39.8)**	**(3.79)**	**(1.36)**
BR2	**46.2**	**2.92**	**1.12**
R^2	0.8590		
Observations	3,485		

Notes:

Replication of Postal Service Model for regular delivery time shown in Table 32 on p. 78 of Postal Service Report. Note that while Table 32 presents an R-squared of 0.8574, the work papers produced by the Postal Service reflects an R-squared of 0.8586, which matches the R-squared produced in my replication. For comparison, I include clustered standard errors. All linear, quadratic, and cross product terms for mail volume are multiplied by 3600 (to convert hours into

seconds). Variables listed in boldface are not statistically different from zero based on clustered standard errors.

Much has been said in the proceeding about model "complexity."[18] I believe that this issue is intertwined with the issue of how much usable data available the special studies provide.

Let me begin by stating clearly that I have no fundamental objections to the use of flexible functional forms, such as the form contained in the Postal Service's Street Time Study Report. They are, as the Postal Service has pointed out, well accepted in empirical economic work.[19] They offer the advantage of being able to approximate any underlying functional form, and it is this flexibility that accounts for the extent to which they are used in empirical research.

This advantage comes at a price, however. As the number of underlying explanatory variables increases, the number of separate coefficients that have to be estimated expands rapidly, and with it, the amount of data required to estimate those coefficients with an acceptable degree of precision.

My use of the term "complexity" in my original report has elicited arguments from the Postal Service over whose model is more "complex." I would concede that the mathematical form of the model presented in my initial report is somewhat more complex than the model put forward by the Postal Service, although neither invokes mathematical forms one wouldn't expect to find in a typical high school curriculum. However, the Postal Service model includes many more estimated coefficients than the model I presented.

Many aspects of the model presented in my original report with which Professor Bradley takes issue are the results of compromises and simplifications made in an effort to pare down the number of parameters being estimated, and improve the precision with which that reduced set of parameters was estimated.

For example, Professor Bradley poses the question of whether the delivery equation should include aggregated volume measures or individual volume measures.[20] He clearly believes that it is important for the model to include individual volume measures, and on this point he and I are in full agreement. As I noted in my original report:

> In its current implementation, the model recognizes just two types of mail: deviation parcels, and other mail, with the former taking a weight of 11.46, and the latter

[18] See, e.g., Initial Neels Report, pp. 2-4; and Bradley Reply Report at p. 15.
[19] See Street Time Study Report, p. 25.
[20] See Bradley Reply Report, p. 14.

taking a weight of one. Comparable weights need to be developed for the other delivery mail steams (e.g., DPS, Cased Mail, FSS, etc.).[21]

That having been said, however, Professor Bradley appears to be more optimistic than I am about the ability of small datasets to provide precise and reliable estimates of differences in marginal cost for a large number of individual mail categories. It was because of concerns such as these that I pointed out in my original report that in a number of cases the marginal costs reported by the Postal Service were statistically indistinguishable.[22] Professor Bradley appears to have taken these observations to heart, and has included in his report the results of tests of a much more comprehensive set of pairwise comparisons. Based on these tests, he concludes that the various types of mail included in his regular delivery model fall into two categories that he labels high cost and low cost. The former includes FSS and collection volume, while the later includes DPS, Cased Mail and Sequenced Mail.[23] He then proceeds to estimate and present the results of a version of his regular delivery model containing just two aggregated measures of mail volume.[24]

The number of aggregated mail volume measures included in this version of Professor Bradley's model is the same as the number of aggregated mail volume measures included in the nonlinear model presented in my original report. This fact perhaps suggests some convergence of opinion regarding the number of distinctions that can be drawn in a statistically reliable way within the confines of the special study dataset. We do, however, aggregate volume measures differently. I included parcels in the model presented in my original report, and for this reason focused on differences in cost intensity between parcels and other types of mail, as this was likely to be the largest such difference to be found among the volume measures I considered.

Professor Bradley also asks whether the constant elasticity model presented in my initial report is acceptable.[25] He points out that in structuring that model I was "imposing *a priori* restrictions on the cost generating process for city carrier street time" for which I provided "neither operational nor economic restrictions."[26] I agree that the model presented in my initial report did impose a priori restrictions. My goal in doing so, however, was to develop a more parsimonious model that still captured the key elements of the city carrier delivery process. Developing accurate measurements of first order effects – the marginal costs associated with the various mail streams – and the second order effects – the tendency for per piece costs to decline with increases in volume – were clearly important goals. Accounting accurately for third order effects – differences in the rate at which per

[21] See Initial Neels Report, p. 27.
[22] See Initial Neels Report, p. 4
[23] See Bradley Reply Report, p. 22.
[24] See Bradley Reply Report, Table 6, p. 23.
[25] See Bradley Reply Report, p. 34.
[26] See Bradley Reply Report, p. 34.

piece costs decline with increases in volume, or subtle differences in density effects across mail streams – was a goal I was willing to sacrifice in order to achieve better estimates of first and second order effects.

The availability of the much larger and richer national Form 3999 dataset obviously alters significantly the tradeoff between model complexity and flexibility on the one hand, and statistical precision on the other. At the time I submitted my initial report I had no way of knowing whether I would be granted access to this resource.

IV. Model Development Based on the National Form 3999 Dataset

In addition to identifying a number of weaknesses in the econometric models supporting Proposal 13, my initial report also proposed an alternative approach to the estimation of volume variabilities for city carrier street time.[27] This approach, if fully implemented, would have several advantages over Proposal 13. First, it would rely upon data regularly collected by the Postal Service in its ordinary course of business, and would thus eliminate the need for special studies that are both costly, and that place artificial constraints on the volume of data that can be used in estimating volume variability. This approach would also focus on the long-run variability in volumes and street time, by changing the nature of the data that are used in estimating the relationships between various volume measures and street time. Finally, my proposed approach would relax several assumptions that Proposal 13 imposes on the data – assumptions that have important costing implications and that, as I showed in my initial report, may not be borne out in the facts. These include the assumption that parcel volumes have no impact on "regular delivery" time, the assumption that non-delivery time has the same variability as delivery time, and the assumption that network travel time is fixed.

My initial report included an illustrative model that followed these principles. The data available to me at the time did not permit a full implementation of the approach outlined in that report, but did show promise. It produced variabilities that were quite different from those implied by Proposal 13. As part of that initial filing, UPS requested and later received access to a confidential version of national Form 3999 data for its external counsel and consultants. To the best of my knowledge, this complete version of national Form 3999 – which contains both route-level time and volume measurements for a single day for all city carrier routes in the country and a zip code field which permits that data to be aggregated for zip-level analysis – represented the best available Postal Service to permit me to move towards a more thorough (though not complete) implementation of the analytical approach outlined above.

[27] See, e.g., Initial Neels Report, pp. 17-18, 26.

In this section, I describe the analyses that I have conducted using the national Form 3999 dataset. I begin by describing the steps I took in cleaning the data and preparing a dataset suitable for analysis, including my treatment of outliers and other problematic data points. I will discuss the choice of functional form, considering briefly the statistical properties of the model presented in my initial report, and evaluating whether the much larger sample size provided by the national Form 3999 dataset is sufficient to support the use of a quadratic flexible functional form. Next, I will provide an overview of the model's specification, as well as a detailed discussion of how I developed imputations for several volume measures that are not contained in the national Form 3999 data but that are nonetheless necessary for costing. Finally, I present the results of the main regression equation.

A. Data Preparation

I used the Nonpublic Form 3999 Material Provided in Accordance with Order No. 2455 in USPS-RM2015-7/NP2 as the starting point for construction of my regression dataset. This is the same as the Route Evaluation Dataset used in Proposal 13 to calculate cost pools, with the addition of zip code and route information. This dataset has 140,457 route level observations that are associated with 10,743 zip codes.

I scanned the data to account for anomalies that appeared to reflect data entry errors. Table 2 lists the obvious errors I found, and the number of routes affected by each type of error. Following the conventions documented by the Postal Service in its Street Time Study Report, I dropped all routes with an evaluation date on Sunday.[28] I also dropped all routes with zero or negative gross street time, and all routes where the sum of reported cased letters, cased flats, DPS, FSS, sequenced, and parcel volumes are zero. Finally, I dropped all routes with zero possible deliveries (delivery points).

Table 2: Data Errors Found in the National Form 3999 Dataset

Data Error	Count of Routes
1: Sunday Evaluation Date	117
2: Zero or Negative Gross Street Time	186
3: Zero Sum of Pieces	178
4: Zero Delivery Points	39

Once I dropped these routes, I sum up reported volumes and times by zip code, and scaled up the raw volumes and times for each zip code for which a route was dropped by multiplying each variable by total delivery points in the zip code divided by remaining delivery points in the zip code after the problematic routes are dropped.

[28] See Street Time Study Report, p. 10.

I used three delivery environment characteristics in my regression; two are calculated directly from the national Form 3999 data while the third is calculated using publicly available data. These variables included (1) the number of delivery points per zip code; (2) the fraction of walking routes per zip code; and (3) the number of miles of local streets by square mile. I classified as walking routes all routes with a delivery mode of Foot, Park and Loop, or Other.

B. Outlier Strategy

At various points in my analysis I noted the existence within the national Form 3999 dataset of volume and time which, while not logically impossible, appeared implausible, and very likely represented errors. For example, the largest street time value in the dataset was more than 100 times the 99th percentile street time value. The accelerated schedule established by the Commission for the completion of this phase of my work did not permit the detailed outlier identification and evaluation procedures described in the Street Time Study Report.[29] To deal efficiently with the problem of potential outliers I used the statistical transformation known as Winsorisation, which typically makes estimators more robust to outliers.[30] This procedure limits the influence of extreme values by setting all outliers to a specified percentile of the data. In the analyses I describe I have "Winsorised" all reported and imputed volume variables and gross street time at or above the 99.5th percentile of that variable. This treatment has the effect of setting values for those variables that exceed the 99.5th percentile to equal the value at the 99.5th percentile.

C. Choice of Functional Form

The substantially larger and richer sample provided by the national Form 3999 dataset appropriately required me to reconsider the tradeoffs between the ability of a model to capture accurately all of the nuances of cost causation, and the statistical precision with which the estimated parameters of that model can be estimated. Does this same specification adequately explain variations in street time within the larger national Form 3999 dataset? Or, are there aspects of cost causation that can be adequately captured and measured only by a more complex model? The answer to this question has critical implications for the choice of functional form.

To answer this question I began by re-estimating the model presented in my initial report on the full national Form 3999 dataset. This regression equation took the following form:

$$ST = \alpha DP^{\gamma_1}\left(\frac{SM}{DP}\right)^{\gamma_2}(NPV + \beta PV)^{\gamma_3}(1 + \delta DM) + \varepsilon \qquad (1)$$

Where:

ST = Gross Street Time

[29] See Street Time Study Report, pp. 80-85 and 111-116.

[30] See Barnett, V. and Lewis, T. 1994. Outliers in Statistical Data. Chichester: John Wiley.

DP = Delivery Points

SM = Street Miles

NPV = Non-parcel volume

PV = Parcel volume

DM = Delivery mode indicator

$\alpha, \beta, \delta, \gamma_1, \gamma_2,$ and γ_3 are parameters to be estimated, and

ε is an error term.

In this regression, I used parcel volume as it is included in the national Form 3999 dataset, recognizing its limitations. I discuss below exactly what that variable measures (or doesn't measure). Non-parcel volumes include all remaining volume variables that are recorded in the national Form 3999 dataset. While this may omit some potential drivers of gross street time, I am constrained by the limits of the data provided to me. I do not believe that the omission of these other potential cost drivers prevents this test from providing useful information regarding the selection of which functional form to use in the remainder of my analysis.

The results I obtained are shown in Table 3. They demonstrated that this model continues to perform well, explaining 98 percent of the total variation in net street time across the 10,660 zip codes contained in the national Form 3999 with useful information. The standard errors are small, and the t-statistics are very large.

Table 3: Results of Non-Linear Regression of Gross Street Time Using National Form 3999 Data

Variable	Estimated Coefficient	Heteroscedasticity-Robust T-statistic
CONSTANT	0.00544	23.84
DP Gamma	0.404	47.45
SMDP Gamma	0.0742	17.95
NPV Gamma	0.595	75.98
PV Beta	11.8	8.55
DM Beta	0.328	39.08
R^2	0.976	
Observations	10,660	

Although this was an impressive performance, that impressiveness does not mean that it cannot be improved upon.

To establish a basis for comparison, I estimated on the same dataset a flexible form quadratic model containing the same underlying set of explanatory variables that enter into equation (1). Unlike the model presented in the Street Time Study Report by the Postal Service, this model includes all possible interaction terms between the underlying volume and delivery environment variables.[31] Results of that regression are shown in Table 4. This model is also able to explain a large fraction of the total variation in net street time. The R-Squared values shown in the two tables are not comparable,[32] and so for comparison purposes it is more appropriate to focus on the sums of squared residuals. The residual sum of squares is 2.75 million for the original non-linear model, and 2.39 for the quadratic model – a reduction of nearly 13 percent. Clearly, in absolute terms, the quadratic model fits the data better, although it uses a much larger set of coefficients to achieve this performance. These coefficients appear to have been estimated with reasonable precision. 28 out of the 45 estimated coefficients take values that differ significantly from zero at the 5 percent level. Three of the estimated coefficients have t statistics in excess of 10, a threshold that none of the coefficients of the Street Time Study Report model achieve once within zip code clustering has been accounted for.[33]

[31] That model included delivery mode and land area per delivery point natural form and squared, but did not include interaction terms between those variables and the volume measures included in the model. No explanation was offered for this omission.

[32] The regression procedure measures the total sum of squares around the mean of the dependent variable, while the non-linear least squares routine measures variation around zero.

[33] Since the dataset underlying the results shown in Table 3 and Table 4 contains only one observation per zip code, no such adjustments are needed here.

Table 4: Results of Flexible Quadratic Form Regression of Gross Street Time Using National Form 3999 Data

Variable	Estimated Coefficient	Heteroscedasticity-Robust T-statistic	Variable	Estimated Coefficient	Heteroscedasticity-Robust T-statistic
DPS	0.000444	2.85	CM	0.00235	9.02
DPS2	(0.0000000186)	(2.43)	CM2	(0.0000000390)	(2.44)
DPS*FSS	0.0000000693	2.33	CM*PAR	0.000000642	1.68
DPS*SEQ	(0.0000000248)	(1.31)	CM*DP	(0.000000125)	(4.55)
DPS*CM	0.0000000403	2.03	CM*DM	(0.00153)	(5.55)
DPS*PAR	0.000000378	1.45	CM*SMDP	0.00235	1.99
DPS*DP	0.0000000801	3.84	PAR	0.0185	3.69
DPS*DM	0.00143	8.85	PAR2	(0.0000177)	(4.61)
DPS*SMDP	(0.00235)	(3.29)	PAR*DP	0.000000530	0.97
FSS	0.00356	8.36	PAR*DM	(0.00514)	(1.05)
FSS2	(0.0000000888)	(1.77)	PAR*SMDP	(0.0593)	(1.33)
FSS*SEQ	0.0000000673	1.59	DP	0.00620	19.62
FSS*CM	(0.0000000725)	(1.63)	DP2	(0.000000152)	(5.87)
FSS*PAR	0.0000000265	0.05	DP*DM	0.000234	0.86
FSS*DP	(0.000000172)	(3.75)	DP*SMDP	0.00240	1.51
FSS*DM	(0.000942)	(2.40)	DM	19.3	11.32
FSS*SMDP	0.0112	1.30	DM2	(15.69)	(10.46)
SEQ	(0.000169)	(0.57)	DM*SMDP	(3.11)	(2.39)
SEQ2	(0.000000140)	(5.25)	SMDP	6.03	5.13
SEQ*CM	0.0000000970	2.98	SMDP2	(0.00822)	(1.78)
SEQ*PAR	(0.00000192)	(3.71)	INTERCEPT	(4.71)	(8.83)
SEQ*DP	0.000000193	5.25			
SEQ*DM	0.000278	0.98			
SEQ*SMDP	(0.00190)	(1.85)			
R^2	0.946				
Observations	10,660				

In choosing between these two model forms it would be helpful to know whether in some sense the differences between them are statistically significant. The fact that the two forms are not nested in the sense normally used in hypothesis testing makes it difficult to answer this question in a rigorous way. However, the fact that the quadratic flexible form is often interpreted as a second order Taylor Series expansion of an unknown function suggests an approximate way forward.

A standard test employed in the evaluation of linear regression models tests the null hypothesis that all of the coefficients in the model are equal to zero, and asks whether that null hypothesis can be rejected. A variant of that test is potentially applicable here. In this case, my null hypothesis would be that the true coefficient values are equal to the values calculated through a hypothetical second order Taylor Series approximation of the nonlinear model described in Table 3. In such a test I would evaluate whether the reduction in the sum of squared residuals associated with moving from the Taylor Series approximation of the nonlinear model to an unconstrained Taylor approximation of a potentially different function indicates that the that null hypothesis can be rejected in favor of the coefficients estimated using the unconstrained quadratic model.

Of course, I did not know the coefficients of this hypothetical approximation of the nonlinear model shown in Table 3, but to implement a version of the test described above I did not need to. Instead, I could calculate the sum of squared residuals directly using the nonlinear model. These residuals differ from those of the hypothetical approximation by the magnitude of the approximation error. Residuals calculated directly from the nonlinear model should be lower, and so using them should reduce the likelihood of rejecting the null hypothesis.

The calculation of the F statistic corresponding to a test of the null hypothesis represented by the nonlinear model shown in Table 3 against the alternative quadratic model shown in Table 4 is shown in Table 5. I divided the reduction in the sum of squared residuals associated with the move from the nonlinear model to the quadratic model by the difference between the number of coefficients in the quadratic model and the number of coefficients in the nonlinear model. I then divided this result by the mean squared error of the quadratic model. The F statistic calculated in this way soundly rejects the null hypothesis that the improvement in fit achieved in moving from the nonlinear to the quadratic model is the result of chance.

Table 5A: : Analysis of Variance Statistics From Tables 3 and 4 Regression Analyses

		Residual Sum of Squares [1]	Degrees of Freedom [2]	Mean Square Error [3]
Quadratic Form				
Model	[a]	42,310,025.50	44	961,591.49
Residual	[b]	2,394,811.68	10,615	225.61
Total	[c]	44,704,837.10	10,659	4,194.09
Non-Linear Form (Parcels and Non-Parcel Pieces)				
Model	[d]	111,084,388.00	6	18,514,064.67
Residual	[e]	2,746,439.17	10,654	257.78
Total	[f]	113,830,827.00	10,660	10,678.31

Notes and Sources:
[1]-[2]: Output of Stata models on gross street time.
[3]: [1] / [2].

Table 5B: Pseudo F Test of the Flexible Quadratic Model Against the Nonlinear Model

		Notes	Value
Total Residual Sum of Squares	[1]	Table 5A, [1][c]	44,704,837.10
Null Hypothesis Residual Sum of Squares	[2]	Table 5A, [1][e]	2,746,439.17
Final Model Residual Sum of Squares	[3]	Table 5A, [1][b]	2,394,811.68
Delta Residual Sum of Squares	[4]	[2] - [3]	351,627.49
Delta Degrees of Freedom	[5]	Table 5A, [2][a] - Table 5A, [2][d]	38
Error Degrees of Freedom	[6]	Table 5A, [2][f] - Table 5A, [2][a]	10,616
Numerator	[7]	[4] / [5]	9,253.35500
Denominator	[8]	[3] / [6]	225.58512
F-Test	[9]	[7] / [8]	41.01935
Probability of Rejection if the Null Hypothesis is True	[10]	F distribution	0.00000

It is possible that the apparent differences between the nonlinear and quadratic forms result simply from the fact that equation (1) aggregates all of the non-parcel mail volumes into a single variable. In effect, this aggregation imposes an a priori restriction that all of the different non-parcel mail streams have the same per piece work content. To account for the possibility that this restriction alone accounts for the differential performance of the two models, I estimate an alternative version of

equation (1) that sets the work content of DPS mail at 1.0, and allows the relative per piece work content of the other mail streams to differ.[34] The specification is as follows:

$$ST = \alpha DP^{\gamma_1}\left(\frac{SM}{DP}\right)^{\gamma_2}\left(DPS + \beta_{CM}CM + \beta_{FSS}FSS + \beta_{SEQ}SEQ + \beta_{PV}PV\right)^{\gamma_3}(1 + \delta DM) + \varepsilon \qquad (2)$$

Where:

ST = Gross Street Time

DP = Delivery Points

SM = Street Miles

DPS = DPS volume

CM = Cased Mail Volume

FSS = FSS Volume

SEQ = SEQ Volume

PV = Parcel volume

DM = Delivery mode indicator

α, β_{CM}, β_{FSS}, β_{SEQ}, β_{PV}, δ, γ_1, γ_2, and γ_3 are parameters to be estimated, and

ε is an error term.

Results of this estimation are shown in Table 6. The estimated coefficients are highly significant. The work content of the different mail streams differs substantially, and these differences are highly significant. The sum of squared residuals equals 2.60 million – a value that is lower than that of my original model, but still well above that of the quadratic model. The coefficients on the separate volume variables are measured with a high degree of precision.

[34] It is necessary to anchor the work content of at least one mail stream in order to avoid introducing perfect multicollinearity into the model.

Table 6: Results of Non-Linear Regression of Gross Street Time with Separate Volume Variables Using National Form 3999 Data

Variable	Estimated Coefficient	Heteroscedasticity-Robust T-statistic
CONSTANT	0.00681	22.74
DP Gamma	0.413	42.19
SMDP Gamma	0.0712	18.76
Pieces Gamma	0.575	63.71
PV Beta	10.2	8.99
DM Beta	0.344	39.95
CM Beta	0.275	5.00
FSS Beta	1.19	10.65
SEQ Beta	0.656	10.36
R^2	0.977	
Observations	10,660	

The calculation of the F statistic corresponding to a test of the null hypothesis represented by the nonlinear model shown in Table 6 against the alternative quadratic model shown in Table 4 is shown in Table 7. While there is some improvement in fit provided by allowing the nonlinear model to estimate separate work content coefficients for the various mail types, I still reject the null hypotheses of the nonlinear model in favor of the quadratic flexible form.

Table 7A: Analysis of Variance Statistics from Tables 3 and 6 Regression Analyses

		Residual Sum of Squares [1]	Degrees of Freedom [2]	Mean Square Error [3]
Quadratic Form				
Model	[a]	42,310,025.50	44	961,591.49
Residual	[b]	2,394,811.68	10,615	225.61
Total	[c]	44,704,837.10	10,659	4,194.09
Non-Linear Form (All Volume Types)				
Model	[d]	111,234,604.00	9	12,359,400.44
Residual	[e]	2,596,222.92	10,651	243.75
Total	[f]	113,830,827.00	10,660	10,678.31

Notes and Sources:
[1]-[2]: Output of Stata models on gross street time.
[3]: [1] / [2].

Table 7B: Pseudo F Test of the Flexible Quadratic Model against the Nonlinear Model with Separate Volume Coefficients

		Notes	Value
Total Residual Sum of Squares	[1]	Table 7A, [1][c]	44,704,837.10
Null Hypothesis Residual Sum of Squares	[2]	Table 7A, [1][e]	2,596,222.92
Final Model Residual Sum of Squares	[3]	Table 7A, [1][b]	2,394,811.68
Delta Residual Sum of Squares	[4]	[2] - [3]	201,411.24
Delta Degrees of Freedom	[5]	Table 7A, [2][a] - Table 7A, [2][d]	35
Error Degrees of Freedom	[6]	Table 7A, [2][f] - Table 7A, [2][a]	10,616
Numerator	[7]	[4] / [5]	5,754.60686
Denominator	[8]	[3] / [6]	225.58512
F-Test	[9]	[7] / [8]	25.50969
Probability of Rejection if the Null Hypothesis is True	[10]	F distribution	0.00000

While I recognize that the tests described above are not exact, I find that they provide compelling evidence that the flexible quadratic form provides a significantly better description of the relationship between gross street time and the available measures of mail volume and the mail delivery environment than the nonlinear model presented in my initial report. The substantially larger estimation dataset provided by the Form 3999 data permits me to estimate the many coefficients of

this model with a high degree of precision. In order, therefore, to provide the most accurate possible estimates of the relationships between mail volume and street time, and to eliminate a source of disagreement and controversy between myself and the Postal Service, I will use the flexible quadratic form for the remainder of my analysis.

D. Model Specification

Based on the tests and evaluations described above, I relied upon a model based on the quadratic flexible form, upon which the model presented in Proposal 13 was based. However, it differs in important ways that should not be ignored by the Commission.

The dependent variable for my analysis is equal to gross street time minus the sum of accountables time and blue box mail collection time, all as reported in the 3999 dataset. Gross street time is the broadest measure of delivery time contained in this dataset, and so was the measure of choice to use in taking a broad view of cost causation. I was unable to develop an acceptable imputation methodology for accountable mail volume, and so elected to treat the corresponding pool of costs, which is relatively small, separately. Similarly, I do not, to the best of my knowledge, have access to a reliable data source for blue box collection volumes or a method for imputing them, whether at a route or zip code level, for the dates in the national Form 3999 dataset.

I related the gross street time dependent variable to seven measures of mail volume (DPS volume, SEQ volume, Cased mail, FSS mail, In-Receptacle Parcels, Deviation Parcels and Collection volume), and three variables describing the delivery environment (number of delivery points, local street miles per delivery point, and percentage of the routes in the zip code that are walk). Three measures of volume (In-Receptacle Parcels, Deviation Parcels, and Collection volume) are not recorded regularly by the Postal Service in the national Form 3999 dataset, and thus needed to be imputed for the purposes of the current analysis.[35] Those imputations are described below, followed by a presentation and discussion of the results from the full regression of gross street time.

My regression equation included as explanatory variables the full set of volume and delivery environment variable in their natural forms, squared and in multiplicative interaction terms with all other variables.

E. Imputation of Special Study Volumes

The parcel volume data provided in the national Form 3999 dataset do not correspond exactly to any of the volume measures used in the Postal Service's city carrier costing analysis. In particular, while Proposal 13 relies on distribution keys for the groupings known as in-receptacle parcels and deviation parcels, the parcels measure in the national Form 3999 dataset represents "a subset of the (generally

[35] See Street Time Study Report, pp. 27, 90.

unknown) sum of in-receptacle and deviation parcels."[36] The Postal Service describes the parcels measure in the Form 3999 data as reflecting a manual count of parcels (originally collected in the DOIS dataset) that are either (a) larger than a shoebox or (b) heavier than two pounds. One would expect most parcels falling into this category to be deviation parcels, but some could be delivered as in-receptacle parcels if the recipient's receptacle is large enough. Thus, positive parcel volume recorded for a given route in the national Form 3999 dataset could represent in-receptacle parcels, deviation parcels, or some combination of the two. At the same time, when zero parcel volumes are recorded for a route in the national Form 3999 dataset, it is still possible that there are nonetheless in-receptacle parcels, deviation parcels, or both, if it were the case that the parcels delivered on a given route for that day do not meet the criteria outlined above.[37] Thus, the national Form 3999 dataset, in and of itself, is not fully informative about the precise volumes of the parcel types addressed by the cost study, and for which distribution keys have been developed.

Similarly, the national Form 3999 dataset lacks measures of collection volume, which is thought to be one of the drivers of city carrier street time costs.[38]

Accordingly, in order to develop an econometric model using the national Form 3999 data while accounting for all relevant cost drivers, I used data from the special studies in order to impute estimated volumes for deviation parcels, in-receptacle parcels, and collection volumes. I employed this strategy, hoping that in the future it will be possible to augment the national Form 3999 dataset to include the missing volumes.

While the national Form 3999 dataset also lacks volume measures for accountables, I decided for several reasons to drop these volumes (or an imputed estimate of these volumes) from the model. First, they make up a very small share of total mail volumes (representing just 0.12% of pieces delivered over the package study period, where pieces is defined as DOIS DPS, FSS, Sequenced, and Cased Mail plus Deviation Study Deviation Parcels and Accountables, plus In-Receptacle Study In-Receptacle Parcels). Secondly, those volumes are small and highly variable and thus difficult to predict with precision; and third, the time spent delivering accountables is separately measured in the

[36] See Notice of the United States Postal Service of Providing Informal Responses to UPS Questions ("USPS Responses to UPS Questions"), Docket RM2015-7, May 28, 2015 at 2: "The parcel counts in the DOIS file are a subset of the (generally unknown) sum of in-receptacle and deviation parcels."

[37] For example, consider the case where two parcels were delivered on a route on a given day. Parcel A weighs less than two pounds, is smaller than a shoebox, and fits inside the receptacle. Parcel B also weighs less than two pounds and is smaller than a shoebox, but does not fit inside the receptacle and is thus counted as a deviation parcel. In this hypothetical case, the 3999 parcels measure would be 0 but would be positive for both IRP and DP.

[38] See Bradley Reply Report, p. 21.

national Form 3999 data.[39] Below I discuss in more detail how the attribution of accountables time can be accomplished.

1. Overall Approach

Postal Service has provided DOIS data for the special study periods, and as the volumes from the national Form 3999 study are taken from the DOIS data, I was able to estimate volume imputation models that can be applied to the full set of zip codes contained in the national Form 3999 dataset. I estimated a relationship between (1) explanatory variables that include measures available in the national Form 3999 dataset (such as volume measures and the number of delivery points, and day of the week), and in some cases, additional zip-code level geographic or socio-economic variables and (2) dependent variables representing the volume measures of interest, which are only available from the special studies for a two-week period. I then applied the resulting equations to the full set of 10,000 plus zip-code level observations in order to impute the missing volume measures. On rare occasions, the imputed values for a given volume type per delivery point will exceed the corresponding maximum observed in the special study. On such occasions, to screen out unreasonable outliers I censored the imputed volume measure, setting it at the corresponding threshold after scaling for the number of delivery points in a zip code. This adjustment affects 314 (of more than 10,000) imputed values for in-receptacle parcels model, 866 imputed deviation parcel volumes, and 182 imputed values for collection volumes.

a. Estimation Method

I estimated negative binomial regression models for each of the three imputation models. The negative binomial is an econometric technique often used to model count data that do not satisfy the poisson assumption that the conditional mean is equal to the conditional variance.[40] The estimating equations take the form:

$$\ln(y) = x\beta$$

where y is the count of volume for a given zip code-day observation, *ln(y)* is its natural log, x is the vector of explanatory variables, and β is the parameter vector. The coefficients have an additive effect on *ln(y)* and thus a multiplicative effect on y.

---

[39] Unlike some of the other subdivisions of time that I criticized as being artificial and thus unreliable, I believe that the discreet incremental activity associated with delivering accountables may be more reliably measured.

[40] When this assumption is not met, the dependent variable is said to be over-dispersed. Standard statistical tests performed on each estimation provide strong evidence that the data are indeed over-dispersed in all three categories. These are provided in my workpapers. I note that the collection volume as reported in the special study does not, strictly speaking, represent count data, as it takes non-integer values. However, these collection volumes are themselves estimates based on linear measurements multiplied by a conversion factor, and these estimates are used by the Postal Service as their piece count for collection volume.

Each of the three models discussed above was based on this general functional form and was estimated on more than 3,000 observations covering the roughly 300 zip codes from one of the two special studies described in the Street Time Study Report. Because observations within a zip code are unlikely to be statistically independent, I estimated standard errors that are clustered at the zip code level. Failure to do so would likely have the effect of overstating the precision of any estimated coefficients.

b. Model Specification

I tuned the general specification described above to the specific needs of each of the three imputation models. In the model selection process, I opted for parsimony whenever possible, since working with the special study data confronted me with the same data constraints I have discussed above and in my initial report. However, in some cases I found that additional socioeconomic variables added sufficient explanatory power to warrant their inclusion.

For deviation parcels, information provided by the Postal Service indicates that the parcels volume measure in the national Form 3999 data is most closely related to deviation parcels, and my preliminary analyses indicated that this is indeed the case.[41] Accordingly, I employ a relatively parsimonious specification in the imputation model for deviation parcels. Specifically, my model predicts deviation parcels (*Devpar*) as a function of delivery points in a zip code-day (*DP*), the average number of DOIS parcels (*Parcels*) per delivery point for that zip code-day observation, the average number of DPS letters (*DPS*) per delivery point, and the average number of cased flats (CF) per delivery point, and an indicator variable for the day of the week (*DOW*) for each observation:

$$\ln(Devpar) = \beta_0 + \beta_1 DP + \beta_2 * \left(\frac{Parcels}{DP}\right) + \beta_3 * \left(\frac{DPS}{DP}\right) + \beta_4 * \left(\frac{CF}{DP}\right) + \sum_{d=2}^{6} \delta_d * DOW_d.$$

Monday (DOW_1) is the reference category for the day of the week and is thus omitted from the specification.

The parcels measure contained in national Form 3999 data is less closely related to in-receptacle parcel volumes than it is to deviation parcel volumes.[42] The model for in-receptacle parcels therefore lacks an anchoring proxy measure like the DOIS parcel counts, and so uses a richer set of explanatory variables, including social and economic information from public sources. I find that once these socioeconomic variables are included in the model, the DOIS parcels measure is not strongly or significantly associated with in-receptacle parcel volumes. My imputation model for in-receptacle parcels (*IRP*) included the total number of delivery points (*DP*), the average number of DPS letters

[41] See USPS Responses to UPS Questions at 2: "In sum, the parcels entered in DOIS are generally the larger/heavier pieces, so they are more likely to be deviation parcels, but they could also be in-receptacle parcels."

[42] See footnote 41.

(*DPS*) per delivery point, and the average number of cased flats per delivery point. I also included measures meant to capture the effects of average household and employer size – the zip code's population (*POP*) divided by the number of delivery points and the number of employees (*EMP*) in the zip code, also divided by delivery points. I also included mean household income (*INC*) and the share of employment in the zip code that are in various sectors of the economy.[43] Specifically, I sum up employment in a zip code according to two-digit NAICS codes, and divide by total employment in the zip code.[44] I again included the day-of-week indicators described above. The final result was the following equation:

$$\ln(IRP) = \beta_0 + \beta_1 DP + \beta_2 * \left(\frac{DPS}{DP}\right) + \beta_3 * \left(\frac{CF}{DP}\right) + \beta_4 * \left(\frac{POP}{DP}\right) + \beta_5 * \left(\frac{EMP}{DP}\right) + \beta_6 * INC + \sum_{s=1}^{18} \gamma_s * NAICS + \sum_{d=2}^{6} \delta_d * DOW_d$$

The omitted NAICS sector is 81: Other Services (except Public Administration).

Finally, for collection volumes (*CV*) I used a slightly different specification. I again included delivery points (*DP*) and average DPS volume per delivery point, the day-of-week indicators, as well as one two-digit NAICS code that appeared particularly influential (56, which encompasses administrative services and support services and thus generates high volumes of mail to be collected). I also included variables indicating the fraction of the population in a zip code that fell into various age ranges (*AGESHARE*), as the amount of mail presented for collection appears to vary with age; specifically I calculate the share of a zip code's population that is under the age of 10, between 10 and 19, and so on. I also included census regional division indicators (*DIVISION*), as there was significant regional heterogeneity in collection mail volumes. Accordingly, I estimated the following equation:

[43] I also merged in zip-code level data on employment by sector, population by age bracket, household income, and Census regional division. In order to gather this data, I have obtained age, population and income data on a zip-code level from the 2013 American Community Survey, via the U.S. Census Bureau. I also used a mapping from U.S. census regions to zip codes. I then merge in 2013 zip code business patterns from the U.S. County Business Patterns database to obtain total employment and employment by sector. Since the zip code business patterns give a range for the number of employees by business type, I assume that the employment for these businesses is the median of the specified range. For example, the Zip Code Business Patterns data might specify that there are 12 companies in a specific sector with 5-10 employees. I assume they each have 7.5 employees in generating the total employment for the sector.

[44] The North American Industry Classification System (NAICS) is used by businesses and governments to classify business establishments according to their industry. At the most detailed level, the system employs a six-digit code, while the most aggregated level uses two-digit codes, which is the level of aggregation that I am using.

$$\ln(CV) = \beta_0 + \beta_1 DP + \beta_2 * \left(\frac{DPS}{DP}\right) + \gamma_1 * NAICS56 + \sum_{d=2}^{6} \delta_d * DOW_d + \sum_{a=1}^{8} \alpha_a * AGESHARE_a + \sum_{r=1}^{8} \rho_r * DIVISION_r$$

2. Results

c. Deviation Parcels

While my imputation model for deviation parcels was very parsimonious, the model still explained 54% of the variation in deviation parcels during the special study. The regression results for this model are presented in Table 8. As the information filed by the Postal Service on May 28 suggests should be the case, the volume of national Form 3999 parcels per delivery point is a strong predictor of deviation parcels, as are the total number of delivery points and the volume of DPS mail. The day-of-week dummies indicate that Mondays and Tuesdays are both lower volume days in terms of deviation parcels.

Table 8: Deviation Parcels Imputation Regression

Variable	Estimated Coefficient	Cluster-Robust T-Statistics
Delivery Points (DP)	0.0000683	16.56
DOIS Parcels per DP	3.90	5.50
DPS Letters per DP	0.148	7.52
Cased Flats per DP	0.0221	0.51
DOW=Tuesday	0.0631	1.38
DOW=Wednesday	0.266	6.74
DOW=Thursday	0.271	6.91
DOW=Friday	0.247	6.69
DOW=Saturday	0.155	4.37
Pseudo-R^2	0.544	
Observations	3,333	

d. In-Receptacle Parcels

The in-receptacle imputation model, whose results are summarized in Table 9 explained roughly 63% of the variation in in-receptacle parcels from the collection study. The coefficients on delivery points, DPS volume per delivery point, and average household income were positive and strongly significant. There was a significant drop off in IRP volumes from Monday to Tuesday, and the day-of-week coefficients were jointly significant. As population and employment per delivery point

increase, in-receptacle parcel volumes decrease, although the coefficient on the latter variable was not statistically significant. Other than a strong negative coefficient on the first NAICS group (Agriculture, Forestry, Fishing, and Hunting), the NAICS coefficients have a limited impact individually, but an F-test confirms that they were nonetheless jointly significant. The coefficient on cased flats – a category that, according to the Postal Service, includes some small parcels – was also insignificant.[45] As I noted above, early exploratory analyses indicated the DOIS parcels measure were a poor predictor of in-receptacle parcels, once socio-economic factors were controlled for.

[45] See Bradley Reply Report, p. 4.

1

Table 9: In-Receptacle Parcel Imputation

Variable	Estimated Coefficient	Cluster-Robust T-Statistics
Delivery Points (DP)	0.0000664	15.05
DPS Letters per DP	0.0672	4.59
Cased Flats per DP	0.0313	1.07
Population per DP	(0.0481)	(2.24)
Employment per DP	(0.0252)	(1.25)
Mean Household Income	0.00000426	7.43
NAICS 11: Agriculture, Forestry, Fishing and Hunting	(14.7)	(2.53)
NAICS 21: Mining, Quarrying, and Oil and Gas Extraction	(2.94)	(1.42)
NAICS 22: Utilities	2.09	1.05
NAICS 23: Construction	0.962	0.71
NAICS 31: Manufacturing	0.426	0.39
NAICS 42: Wholesale Trade	1.75	1.59
NAICS 44: Retail Trade	1.15	1.10
NAICS 48: Transportation and Warehousing	(0.114)	(0.09)
NAICS 51: Information	(0.0924)	(0.07)
NAICS 52: Finance and Insurance	(0.287)	(0.27)
NAICS 53: Real Estate and Rental and Leasing	(2.19)	(0.83)
NAICS 54: Professional, Scientific, and Technical Services	0.958	0.90
NAICS 55: Management of Companies and Enterprises	(0.111)	(0.09)
NAICS 56: Administrative and Support	1.48	1.35
NAICS 61: Educational Services	(0.0354)	(0.03)
NAICS 62: Health Care and Social Assistance	0.770	0.70
NAICS 71: Arts, Entertainment, and Recreation	0.879	0.74
NAICS 72: Accommodation and Food Services	0.224	0.20
NAICS 99: Unclassified	(10.4)	(0.25)
DOW=Tuesday	(0.344)	(8.32)
DOW=Wednesday	(0.0658)	(2.03)
DOW=Thursday	(0.00170)	(0.05)
DOW=Friday	0.00690	0.22
DOW=Saturday	(0.0857)	(3.01)
Pseudo-R^2	0.626	
Observations	3,333	

2

e. Collection Volume

The collection volume results are summarized in Table 10. This model explained roughly 32% of the variation in collection volumes during the special study.[46] I found that collection volumes were closely associated with the number of delivery points and with DPS volumes. The large and significant coefficient on *NAICS56* showed that the sector does in fact generate large amounts of collection volume. Collection volumes tended to be higher on mid-week days (Tuesday through Friday) than on Mondays and significantly lower on Saturdays. Census regional divisions 3, 4, 6, and 8 (corresponding to parts of the Midwest, South, and West) generated high levels of collection volume, with all significantly higher than in New England, the reference division. Collection volumes also varied by age of the population in that zip code; one notable result among these coefficients was that the collection volume increases with the share of a zip code's population that is 80 or older. While not all of the age and Census Division coefficients were individually significant, an F-test run on either group of coefficients strongly rejects the null hypothesis that these variables have no impact on the estimation.

[46] This pseudo-R^2 measure is based on comparing the predicted values from the negative binomial regression with the actual values.

Table 10: Results of Imputation Model for Collection Volume

Variable	Estimated Coefficient	Cluster-Robust T-Statistics
Delivery Points (DP)	0.0000434	6.52
DPS Letters per DP	0.153	5.41
NAICS 56: Administrative and Support	2.89	4.09
DOW=Tuesday	0.192	3.79
DOW=Wednesday	0.216	4.45
DOW=Thursday	0.176	4.30
DOW=Friday	0.172	4.08
DOW=Saturday	(0.264)	(5.40)
Age Group: Under 10	(0.0285)	(2.21)
Age Group: Between 10 and 19	(0.0259)	(1.94)
Age Group: Between 30 and 39	(0.0443)	(2.21)
Age Group: Between 40 and 49	0.0155	0.91
Age Group: Between 50 and 59	0.0168	0.80
Age Group: Between 60 and 69	(0.0724)	(3.37)
Age Group: Between 70 and 79	(0.0529)	(1.74)
Age Group: Over 80	0.0892	2.87
Census Division 2: Middle Atlantic	0.125	0.59
Census Division 3: East North Central	0.419	2.03
Census Division 4: West North Central	0.688	2.86
Census Division 5: South Atlantic	0.220	1.07
Census Division 6: East South Central	0.593	2.38
Census Division 7: West South Central	0.342	1.61
Census Division 8: Mountain	0.789	3.74
Census Division 9: Pacific	0.218	1.09
Pseudo-R^2	0.315	
Observations	3,501	

f. Accountables

I was not able to develop an acceptable imputation method for accountables volume. Although I developed some specifications that appeared plausible upon first examination, in using them to impute accountables volumes across the entire Form 3999 dataset I found that they tended to produce extreme outliers, and never yielded reasonable looking results when used in the street time model. In retrospect, this outcome is perhaps not surprising. Accountables volumes are small, and likely to be heavily influenced by idiosyncratic factors.

Below I discuss in more detail how I think accountables ought to be handled if the Commission should decide to adopt the results of this analysis for the costing of the other mail streams.

g. Imputation

The final step in the imputation process was to use the estimated coefficients to calculate the missing volumes for all 10,000 zip codes in the national Form 3999 dataset. As mentioned above, I censored the predicted values such that the imputed volume per delivery point for a given type of volume cannot exceed the largest observed volume per delivery point in the special study for that stream of mail. Note that for many of the zip codes in the 3999 dataset, the routes in that zip code were evaluated on different days of the week. In such cases, the coefficients associated with day-of-week indicators were applied according to the proportion of total delivery points in a zip code whose evaluations occurred on that day.

F. ESTIMATION

I estimate the coefficients of the full model using ordinary least squares. The regression results from estimation of this model are presented in Table 11 below.

The model explains 95 percent of the variation in net street time. The reported results indicate that of the 66 coefficients (including the constant term) that appear in the model, 42 are statistically significant at the 10 percent level or higher, 36 are significant at the 5 percent level or higher and 30 are significant at the 1 percent level or higher. All of these results are based upon heteroscedasticity-corrected standard errors.

Four of the seven first order terms associated with mail volume measures (FSS, cased mail, deviation parcels and in-receptacle parcels) are statistically significant. In a series of tests of the null hypothesis that the coefficients of all of the terms associated with a specific mail volume measure are equal to zero, the null hypothesis is strongly rejected for all of the mail volume measures. Clearly, all of the mail volume measures belong in the model.

Table 11: 3999 Model Regression Results

Variable	Estimated Coefficient	Heteroscedasticity-Consistent T-statistic	Variable	Estimated Coefficient	Heteroscedasticity-Consistent T-statistic
DPS	(0.000250)	(1.26)	CM*CV	(0.0000000760)	(0.98)
DPS2	(0.0000000324)	(3.61)	CM*IRP	0.000000823	0.89
DPS*FSS	0.0000000923	2.89	CM*DP	(0.000000180)	(5.30)
DPS*SEQ	(0.0000000153)	(0.73)	CM*DM	(0.00138)	(5.60)
DPS*CM	0.0000000491	2.48	CM*SMDP	0.00153	1.37
DPS*DEVPAR	0.00000217	1.48	DEVPAR	0.123	5.81
DPS*CV	0.0000000947	1.78	DEVPAR2	(0.000203)	(2.94)
DPS*IRP	(0.000000139)	(0.20)	DEVPAR*CV	(0.00000805)	(1.55)
DPS*DP	0.000000090	3.31	DEVPAR*IRP	0.000112	1.56
DPS*DM	0.00156	9.48	DEVPAR*DP	0.00000383	1.37
DPS*SMDP	(0.00202)	(2.68)	DEVPAR*DM	(0.0972)	(5.58)
FSS	0.00356	7.58	DEVPAR*SMDP	(0.0730)	(1.83)
FSS2	(0.000000111)	(2.19)	CV	0.000172	0.25
FSS*SEQ	0.0000000261	0.60	CV2	(0.000000225)	(2.65)
FSS*CM	(0.0000000552)	(1.26)	CV*IRP	(0.00000116)	(0.44)
FSS*DEVPAR	(0.00000273)	(0.93)	CV*DP	0.000000526	6.09
FSS*CV	(0.000000305)	(2.69)	CV*DM	0.00129	2.93
FSS*IRP	0.00000343	2.37	CV*SMDP	(0.000648)	(0.71)
FSS*DP	(0.000000200)	(3.28)	IRP	0.0389	4.46
FSS*DM	(0.000659)	(1.77)	IRP2	(0.0000978)	(4.25)
FSS*SMDP	0.0103	1.19	IRP*DP	0.000000686	0.59
SEQ	0.000542	1.78	IRP*DM	0.00359	0.46
SEQ2	(0.000000151)	(5.57)	IRP*SMDP	(0.0325)	(2.27)
SEQ*CM	0.0000000907	3.01	DP	0.00479	12.13
SEQ*DEVPAR	(0.00000513)	(1.77)	DP2	(0.000000273)	(5.83)
SEQ*CV	(0.000000333)	(3.99)	DP*DM	0.00127	5.02
SEQ*IRP	0.00000223	1.66	DP*SMDP	0.00336	1.99
SEQ*DP	0.000000223	4.44	DM	21.4	9.69
SEQ*DM	0.000277	1.02	DM2	(14.0)	(9.46)
SEQ*SMDP	(0.00157)	(1.32)	DM*SMDP	(0.100)	(0.06)
CM	0.00200	6.58	SMDP	15.5	5.50
CM2	(0.0000000304)	(2.04)	SMDP2	(0.0371)	(4.18)
CM*DEVPAR	0.00000150	0.81	INTERCEPT	(16.1)	(8.34)

These statistical tests do not account for the fact that there is some imprecision in a number of the volume variables. The problems with the DOIS parcel volumes have been extensively discussed. Those problems introduce some noise into the deviation parcel volume variable used in the model. In addition, because the three imputed volume measures are themselves estimates based on first-stage regression estimates they contain some additional error that has not been fully accounted for. However, these problems would be eliminated by data collection improvements that I discuss and recommend below.

V. Results

Table 12 below presents the marginal costs by mail stream measured in terms of incremental net street time per mail piece. It also presents the cost variabilities associated with each mail stream. The point estimates of the marginal costs for DPS and cased mail are very similar. The point estimates of the marginal costs for FSS and collection volume are noticeably higher. This pattern is similar to that found in the Street Time Study Report by the Postal Service. These results imply noticeably lower marginal cost for sequenced mail than the Street Time Study Report. The cost intensities of in-receptacle parcels and deviation parcels are much larger than that of the other 5 types of mail, which is consistent with the notion that these are significantly more cost-intensive than most of the volume delivered by the Postal Service.

Table 12 also shows the standard errors of estimated marginal costs. Based on the results presented above in Table 11, all are measured with a high degree of statistically reliability. However, as noted above, caution is called for in interpreting these statistical results, as they do not account for the statistical uncertainty in the imputed volume measures.

The individual volume variabilities also suggest that the two types of parcels are responsible for 15.9% of total city carrier street costs (less accountables and collection box time).

Table 12: Marginal Costs and Variabilities, by Mail Type, from the Form 3999 Regular Delivery Equation

Mail Category	Marginal Cost (Seconds) [A]	Marginal Cost Standard Error [B]	Variability [C]
DPS	3.60	0.29	21.4%
Cased Mail	3.89	0.42	7.8%
Sequenced	1.41	0.42	1.2%
FSS	12.99	3.09	3.8%
Collection	5.74	0.65	5.4%
In Receptacle Parcels	38.42	10.49	4.1%
Deviation Parcels	162.08	34.25	11.8%

Notes:

[A]: Calculate as the partial derivative of the 3999 Model Equation with respect to the mail category, evaluate at the sample means and converted into seconds.

[B]: Standard Error of Marginal Costs in [A]

[C]: Calculate as marginal cost times the ratio of the sample mean of mail category volume to predicted delivery time at sample means.

VI. Recommendations for moving forward

In the Order modifying the schedule for this docket and providing the opportunity to file this report, the Commission expressed a strong desire to conclude this proceeding in time to make an informed decision regarding the costing and analytical procedures that should be used in the 2015 Annual Compliance Determination.[47] In this section I discuss what I believe are the relevant alternatives for the Commission to consider, and I outline what I believe to be the most appropriate course of action for moving forward.

I divide my recommendations into two parts. I first consider the alternatives that are realistically available in connection with ACD 2015. In my view all of the available alternatives are subject to important limitations. For this reason, following an evaluation of these alternatives, I discuss actions and decisions that I think might improve the options available to the Commission in future years.

A. Costing Options for FY2015

In my view there are four options for the treatment of city carrier delivery cost in FY2015. These are:

- Continued use of legacy costing procedures;
- Adoption of Proposal Thirteen put forward by the Postal Service;
- Adoption of a modified version of Proposal Thirteen that uses econometric results presented in my initial report to assign a percentage of regular delivery costs to parcels; or
- Adoption of the results of the analysis of the national Form 3999 dataset described above.

Table 13 below shows the cost attribution for each of these costing options.[48]

[47] Order No. 2455 - Order Granting United Parcel Service, Inc. Motion for Issuance of Commission Information Request No. 1 and Revising Procedural Schedule, Docket RM2015-7, April 23, 2015.

[48] In calculating the cost implications of adoption of Proposal Thirteen I use the volume variabilities presented in the Street Time Study Report, despite having some reservations regarding their accuracy. That report indicates that the variability of deviation packages is 31.1 percent. See Street Time Study Report, Table 50. However, examination of the library reference accompanying the report -- RM2015.7.1 "Deviation_Acct_Variabilities_Model.sas" code -- suggests that this deviation parcel variability may have been calculated incorrectly. The apparent error involves multiplication by an extraneous dummy variable. Correcting for this error, I arrive at a variability of 33.9 percent, which would imply that the figures shown in Table 13 may understate the attribution for deviation parcels.

Table 13: Comparison of 2013 CS6&7 Cost Attributions by Class for Four Costing Option ($000s)

	ACR 2013- Legacy Costing	Modified Proposal 13	Proposal 13	3999 Model
FIRST-CLASS MAIL				
SINGLE-PIECE LETTERS	1,539,729	1,276,415	1,261,981	1,462,065
SINGLE-PIECE CARDS	84,929	72,077	70,987	82,488
PRESORT LETTERS	1,126,093	1,054,429	1,018,711	1,362,037
PRESORT CARDS	54,670	51,503	49,428	67,024
FLATS	255,214	238,808	233,017	255,560
PARCELS	44,015	51,582	57,638	95,276
TOTAL FIRST-CLASS	3,104,651	2,744,813	2,691,761	3,324,450
STANDARD MAIL				
HIGH DENSITY & SATURATION LETTERS	131,999	145,850	138,465	167,024
HIGH DENSITY & SATURATION FLATS & PARCELS	280,815	462,678	437,786	322,282
EVERY DOOR DIRECT MAIL - RETAIL	23,446	36,850	34,981	23,968
CARRIER ROUTE	588,339	651,954	613,653	767,296
LETTERS	1,381,078	1,282,094	1,235,754	1,659,632
FLATS	529,267	562,314	540,200	627,949
PARCELS	15,114	18,520	20,830	34,429
TOTAL STANDARD MAIL	2,950,059	3,160,261	3,021,669	3,602,579
PERIODICALS				
IN-COUNTY	27,787	30,577	28,529	36,748
OUTSIDE COUNTY	395,317	421,853	402,316	480,328
TOTAL PERIODICALS	423,104	452,430	430,845	517,076
PACKAGE SERVICES				
SINGLE-PIECE PARCEL POST	9,504	9,730	14,618	30,082
BOUND PRINTED MATTER FLATS	19,568	19,284	18,463	22,074
BOUND PRINTED MATTER PARCELS	43,698	47,280	68,895	144,544
MEDIA AND LIBRARY MAIL	20,756	22,698	31,334	62,003
TOTAL PACKAGE SERVICES	93,526	98,993	133,311	258,702
US POSTAL SERVICE	53,054	52,733	53,392	63,237
FREE MAIL	6,042	7,053	7,478	10,592
Total Domestic Market Dominant Mail	6,630,436	6,516,283	6,338,455	7,776,637
Ancillary Services				
CERTIFIED	117,738	93,241	93,241	110,068
COD	523	427	427	493
INSURANCE	6,191	4,351	4,351	4,880
REGISTRY	1,326	1,149	1,149	1,270
OTHER ANCILLARY SERVICES	172,547	48,687	48,687	59,498
TOTAL SPECIAL SERVICES	298,325	147,856	147,856	176,208
Total Domestic Market Dominant Mail and Services	6,928,761	6,664,139	6,486,311	7,952,845
Competitive Products			-	
Total Domestic Competitive Costs	599,424	680,351	925,215	1,819,609
INTERNATIONAL MAIL	57,299	57,539	64,038	99,297
TOTAL VOLUME VARIABLE COSTS	7,585,485	7,402,029	7,475,564	9,871,751
OTHER	8,048,193	8,231,651	8,158,116	5,761,927
GRAND TOTAL	15,633,678	15,633,680	15,633,680	15,633,678
Competitive % of Volume Variable Costs	7.9%	9.2%	12.4%	18.4%
Competitive % of Total Costs	3.8%	4.4%	5.9%	11.6%
Variability	48.5%	47.3%	47.8%	63.1%

Notes: The "ACR2013" and "Proposal 13" columns are taken from the 'Letter Route Study Cost Impact' tab from the 'Cost_Impacts_Proposal_13.xlsx' file contained in the RM2015.7.NP1 library reference. Note that Table 52 of the Report on the City Carrier Street Time Study exhibits slight unexplained differences from this table. The maximum difference for classes reported in both is less than 0.15%.

To demonstrate the difference in these costing options, Table 14 below shows the change in cost attribution from the Legacy ACR 2013 model to each of the three other options.

Table 14: Change in FY13 Cost Attribution by Class for New Costing Options Relative to Legacy ACR 2013 Alternative ($000s)

	Proposal 13	Modified Proposal 13	3999 Model
FIRST-CLASS MAIL			
SINGLE-PIECE LETTERS	(263,315)	(277,749)	(77,664)
SINGLE-PIECE CARDS	(12,852)	(13,942)	(2,441)
PRESORT LETTERS	(71,664)	(107,382)	235,944
PRESORT CARDS	(3,167)	(5,242)	12,354
FLATS	(16,406)	(22,197)	346
PARCELS	7,567	13,623	51,261
TOTAL FIRST-CLASS	(359,838)	(412,890)	219,799
STANDARD MAIL			
HIGH DENSITY & SATURATION LETTERS	13,851	6,465	35,025
HIGH DENSITY & SATURATION FLATS & PARCELS	181,862	156,971	41,466
EVERY DOOR DIRECT MAIL - RETAIL	13,404	11,535	522
CARRIER ROUTE	63,615	25,314	178,957
LETTERS	(98,983)	(145,324)	278,554
FLATS	33,047	10,933	98,682
PARCELS	3,406	5,716	19,314
TOTAL STANDARD MAIL	210,202	71,609	652,520
PERIODICALS			
IN-COUNTY	2,790	742	8,961
OUTSIDE COUNTY	26,537	6,999	85,011
TOTAL PERIODICALS	29,326	7,741	93,972
PACKAGE SERVICES			
SINGLE-PIECE PARCEL POST	226	5,114	20,577
BOUND PRINTED MATTER FLATS	(283)	(1,104)	2,507
BOUND PRINTED MATTER PARCELS	3,582	25,197	100,846
MEDIA AND LIBRARY MAIL	1,942	10,578	41,246
TOTAL PACKAGE SERVICES	5,467	39,784	165,176
US POSTAL SERVICE	(321)	338	10,183
FREE MAIL	1,011	1,436	4,550
Total Domestic Market Dominant Mail	(114,153)	(291,981)	1,146,201
Ancillary Services			
CERTIFIED	(24,497)	(24,497)	(7,671)
COD	(96)	(96)	(30)
INSURANCE	(1,840)	(1,840)	(1,312)
REGISTRY	(176)	(176)	(56)
OTHER ANCILLARY SERVICES	(123,860)	(123,860)	(113,049)
TOTAL SPECIAL SERVICES	(150,469)	(150,469)	(122,117)
Total Domestic Market Dominant Mail and Services	(264,622)	(442,450)	1,024,084
Competitive Products			
Total Domestic Competitive Costs	80,927	325,791	1,220,185
INTERNATIONAL MAIL	239	6,738	41,998
TOTAL VOLUME VARIABLE COSTS	(183,456)	(109,921)	2,286,266
OTHER	183,458	109,923	(2,286,266)
GRAND TOTAL	-	-	-

I discuss each of these in turn.

There is little to be said about the first option. Maintenance of legacy attribution methodology and parameters may be appropriate if the Commission lacks confidence that any of the other alternatives represent an improvement over current practice.

I have little to add at this point to what has already been said about Proposal 13. The conceptual framework and assumptions upon which it rests are similar to those underlying the legacy approach. In my view, therefore, it shares the same flaws that exist in the legacy approach. I don't believe the two adequately capture the effects of growing parcel volumes on the overall delivery process, and I don't believe the key costing parameters are based upon adequately sized samples. Proposal Thirteen, however, has the advantage of being based on much more up to date information and larger and cleaner samples than those underlying legacy costing procedures.

The third option requires some explanation. In my initial report I presented the results of a regression analysis based upon a version of the regular delivery time model that included DOIS parcel counts, entered in natural form, squares and in multiplicative interaction terms with delivery points and all of the other volume variables included in the model.[49] The results of that analysis implied that 2.9 percent of regular delivery time can be attributed to DOIS parcels, which as previously discussed are primarily deviation parcels.[50] Given measurement issues associated with the DOIS parcel counts that have been identified by the Postal Service in this docket, I believe that 2.9% figure likely understates the true share of regular delivery time that parcel volumes account for.[51]

This result provides a straightforward way to address what I believe to be the most serious shortcoming of the Proposal 13 approach – namely the under-attribution of regular delivery time to parcels. My third option, therefore, would be to adopt a modified version of Proposal Thirteen that involves moving 2.9 percent of the regular delivery cost pool to the deviation parcel cost pool and making the admittedly slight adjustments that those results imply with respect to the variabilities for other mail streams, but otherwise following the details of Proposal Thirteen.

The fourth option available to the Commission is to adopt the model presented above that is based on the nationwide Form 3999 dataset. Admittedly, these results have their limitations. They rely upon imputed volume measures for a couple of important mail streams. While these imputed measures appear reasonable and appear to perform well, they are imperfect substitutes for real data.

---

[49] See Initial Neels Report, p. 9.

[50] See Initial Neels Report, p. 10, Table 3.

[51] Professor Bradley has opined that this effect has more to do with a systematic bias in the measurement error in the DOIS parcels measure than with the effect of parcel volumes on regular delivery time. I do not believe that he has sufficiently substantiated this claim, and I maintain that any measurement error is likely to result in attenuation bias (understating the true effect). See Initial Neels Report, p. 8 and Bradley Reply Report, p. 14.

However, this model also has a number of important advantages over the other available options. It considers all city carrier street time in a comprehensive and consistent fashion, and dispenses with the necessity of accepting a number of long held, untested, and increasingly questionable assumptions about the cost structure and cost causation. It is based on a much larger and richer set of information than the alternative proposals. Finally, because it is based on readily available operational data, the entire analysis can be updated next year, if desired, at a modest cost.

The formation of the street time cost pools applicable to the Form 3999 Model is slightly different than the formation of cost pools in Proposal 13. First, Proposal 13 separates Directly Attributable, Indirectly Attributable, and Office Time Costs. Cost Pools are then formed from the Directly Attributable Costs. Table 15 below shows how Directly Attributable Street Time costs are split between cost pools in Proposal 13.[52] The Regular Delivery cost pool includes costs for Cased Mail, DPS, FSS, Sequenced, and Customer Collections. Proposal 13 describes all other times as "Allied Time" which are dealt with separately. Allied time includes delivery and collection activities such as In-Receptacle Parcel Delivery, Parcel/Accountable Delivery, and General and Express Collections. In addition, Allied time includes Network Travel which is entirely fixed and Travel To/From Route and Relay which is allocated between delivery and network travel.

Table 15: Proposal 13 Street Time Cost Pools

Cost Pool	Percent of Street Time
Regular Delivery	78.23%
IR Parcel Delivery	4.40%
Parcel/Accountable Delivery and Travel	5.39%
General Collections	0.20%
Express Collections	0.00%
Travel To/From Route and Relay	8.86%
Network Travel	2.93%

Source: Tab 7.0.4.1 in CS06&7_Proposal_13_NP.xls.

In contrast, adoption of the 3999 Model would require that street time be divided among three major cost pools: Delivery, Accountables, and Box Collections. My Delivery Cost Pool includes delivery of Cased Mail, DPS, FSS, Sequenced, In-receptacle Parcels, and Deviation Parcels as well as Customer Collections. In contrast to Proposal 13, it is not necessary to remove what Postal Service classifies as "Indirectly Attributable" costs from this pool. The Accountables and Box Collections Cost Pools are calculated based on the proportion of Accountable Hours in the national Form 3999 dataset to Gross

[52] Special Purpose Routes and In-Office costs are dealt with separately.

Street Hours. Box Collections is then further split between Priority Mail Express (PME) Collections and General Collections based on PME boxes comprising 0.19% of collection boxes.

Table 16: Form 3999 Model Street Time Cost Pools

Cost Pool		Total Time From 3999 Regression Dataset (Hours)	Percent of Total Street Time
Delivery	[1]	840,941	98.49%
Accountables	[2]	11,463	1.34%
Box Collections	[3]	1,391	0.16%
Express Collections	[4]	3	0.00%
General Collections	[5]	1,388	0.16%
Total Street Time	[6]	853,796	100.00%

Notes and Sources:
[1]: Total of Winsorized Gross Street Time less Accountable Time and Box Collection Time from the 3999 Regression Dataset.
[2]: Total Winsorized Accountable Time from the 3999 Regression Data Set.
[3]: Total Winsorized Box Collection Time from the 3999 Regression Data Set.
[4]: 0.19% of [3] based on Tab 'I-CS 6&7 FACTORS NEW' of I_FORMS_Proposal_13.xls provided in RM2015.7.1, 0.19% of collection boxes that are Priority Mail Express.
[5]: [3] - [4].
[6]: [1] + [2] + [3].

The results derived from analysis of the national Form 3999 dataset do not cover accountables, and so if the national Form 3999 results are to be used, an alternative approach for handling accountables must be devised. In the national Form 3999 Model presented here, I treat accountables as 100% variable. That approach will necessarily be constrained by the available data, and should reflect their characteristics and the importance of this mail stream. As the Postal Service notes, although accountables are costly to handle on a per piece basis, their volume is so low that they account for only a small and relatively stable share of delivery costs.[53] For this reason it is important for the Commission to think carefully about what level of resources and review time it makes sense to devote to this relatively small element of city carrier operations, and about how best to make tradeoffs between improving the accuracy of cost attribution for this activity and improving cost attribution for the larger and more rapidly growing mail streams.

The national Form 3999 dataset contains a measure of accountables delivery time. Although in principal the same arguments made above about parcels also apply to accountables–namely, that they

[53] See Street Time Study Report, p. 86.

constitute an element of an integrated operation, and so are likely to influence overall costs in ways that extend beyond the act of turning individual pieces over to the recipient, in practice their implications for the accuracy of cost attribution are likely to be far more limited, simply because of the much lower volumes involved. For this reason it seems reasonable to me to base cost attribution for accountables on the time associated with them in the national Form 3999 dataset.

There remains the question of determining the portion of accountables time that is variable, and hence that should be attributed to the products that comprise this mail stream. A plausible argument can be made for attributing all of this time to products. If this category includes just the time spent carrying an accountables piece to the recipient and carrying out the necessary interactions with the recipient, it is hard to imagine what portion of this time would remain if the piece were to go away. The only circumstance in which it might be reasonable to argue that a portion of this cost might be fixed would be if an individual recipient were to receive two accountables or an accountable and a deviation parcel on the same day – a situation which, given the volumes involved, would be expected to occur very infrequently. Moreover, to the extent that full attribution of accountables time might result in some small over-attribution of time that ought properly to be regarded as institutional, this effect would be offset in whole or in part by the fact that whatever effects the presence of accountables has on other components of street time are being ignored.

For all of these reasons, I recommend that the accountables time recorded in the national Form 3999 dataset be used to form the accountables cost pool, and that this pool should be fully attributed to products using the existing accountables distribution key.

B. RECOMMENDATIONS

In my view the advantages of the national Form 3999 model strongly outweigh its limitations, and for the reasons discussed in this Supplemental Report, I urge the Commission to adopt the new cost analysis based upon the national Form 3999 dataset for use in measuring the variability of city carrier delivery costs, and in assigning those costs to products. I further believe it is important to establish protocols for assuring that this analysis is routinely updated as new and better data become available. I discuss what I believe to be the key elements of this process below.

I have concluded that the national Form 3999 model can be used to provide accurate and reliable measurements of the effects of changing mail volumes on city delivery costs in a form suitable for ongoing use in product level cost determination. As discussed above, the results of the national Form 3999 model offer many significant advantages over the proposal put forward by the Postal Service. The primary limitation of the national Form 3999 model results, at the present time, is the need to rely on constructed measures of volume for three mail streams – customer collection volumes, deviation parcels, and in-receptacle parcels. This need arises because of limitations in the data collected by the Postal Service for the Form 3999 dataset. For the following reasons, I do not believe

that this current need to rely on constructed measures of volume for these mail streams is a significant problem that undermines the national Form 3999 model. In order to explain this conclusion, I will discuss the roles that these mail streams play, and the implications of those roles for the nature and degree of error in cost attribution that the Commission and the mailing community should be willing to tolerate.

Collection volumes are large in absolute terms, and according to both my results and those of the Postal Service, relatively costly on a per piece basis. However, inspection of the distribution keys for this mail stream indicates that it is dominated by first class mail.[54] As a market dominant product, first class mail is subject under PAEA to a rate cap rather than a cost floor. In addition, first class mail carries a high markup, and contributes disproportionately to institutional costs. Both of these facts suggest that the risks of tolerating some imprecision in the attribution of collection volume costs are relatively low. Modest over or under attribution of costs are unlikely to alter the extent to which first class mail complies with the requirements of PAEA, or force inappropriate changes in rate levels. Thus, the current need to construct volume measures for customer collection models is not, in my view, a significant limitation.

The deviation parcel volumes used in my model are based upon an imputation regression. That regression relies upon a proxy measure of deviation parcel volumes – namely, the DOIS parcel volume counts – that is widely available and that is closely related both conceptually and empirically to deviation parcels. The availability of this measure and its close empirical relationship to the volume measure being estimated gives me confidence in the robustness of the imputation results for deviation parcels, despite the known problems that infect this measure.

In-receptacle parcels are relatively costly on a per piece basis and are growing rapidly in volume. Because this mail stream is heavily skewed toward competitive products, the volume of costs attributed to it plays a role in setting the required cost floor for competitive projects. The imputation model I have developed lacks an "anchor" explanatory variable analogous to the DOIS parcel count variable.[55] To the extent that there are errors in the imputed IRP volumes, those errors could have some effect on costing. The imputation model I have developed for this mail stream, however, appears both conceptually reasonable and statistically reliable.

---

[54] The distribution key for customer collection contained in the "CS06&7_Proposal_13.xls" file contained in the USPS Library Reference accompanying Proposal 13 indicates that First-Class Mail accounts for 90.4% of collection volumes. While there are some minor adjustments made before the $491 million attributed to collection volumes in Proposal 13 is distributed to products, the First-Class share is very similar, at 89.9%.

[55] I experimented with adding the DOIS Parcel variable to the IRP imputation variable, and found that it failed to achieve statistical significance.

Thus on balance, the strengths of the national Form 3999 model strongly outweigh its limitations. I also note that these limitations are not inherent in the model and can be addressed by improved data collection going forward. This adaptability of the model is one of its strengths and one that I thought about as I developed the model. This feature of the model will allow it be improved as additional data is collected in the future. I discuss below the additional data that could be collected to address the current need to construct volume measures.

If the Commission concludes that this approach is not yet ripe for use, I would recommend adoption of the modified version of Proposal Thirteen described above that moves a portion of regular delivery time to the large parcel cost pool. This modification is based on solid statistical evidence, and addresses a significant shortcoming of Proposal Thirteen as presented in the December 2014 report. I see no reason to consider adoption of Proposal Thirteen as presented. Either of the two options discussed above offer significant advantages relative to Proposal Thirteen. Similarly, I see no benefit from continued reliance on legacy costing methods.

C. COSTING OPTIONS GOING FORWARD

If the Commission chooses to rely upon the national Form 3999 model presented above for the attribution of city carrier delivery costs in FY2015, it would make sense for the Postal Service to improve the volume data upon which it is based in order to increase its accuracy and reliability in future years. Even if the Commission decides that these results are insufficiently reliable for use in the current fiscal year, I urge the Commission to require the Postal Service to address the data gaps that limit the utility of the national Form 3999 data for costing purposes. In either case, I believe it is important for the Commission to require the Postal Service to improve the quality of the volume data that it collects.

In an ideal world it would be desirable for the Postal Service to collect as part of its regular route evaluation process reliable volume data for all of the mail streams for which costs need to be attributed. This would mean augmenting the national Form 3999 dataset to include volume measures for deviation parcels, in-receptacle parcels, accountables, customer collection volume, and mail collected from street-letter boxes. The existence of the national Form 3999 process, which evaluates every city carrier route in the country on a regular basis, provides a natural opportunity to circumvent the need for special studies.[56] It also has added advantages: first, the route evaluations are conducted throughout the year and thus, if correctly implemented, has the potential to incorporate seasonal variation into variability estimates. In addition, retention of data from past route evaluations would allow for the establishment of longitudinal variation over time that is more germane to the

[56] See, e.g. City Carrier Study Report, p. 5. 99.5% of all routes were evaluated at least once in a three-and-a-half year period, and 96% of all routes had been evaluated at least once in a two-and-a-half year period.

concept of cost causation that I believe the Commission is interested in measuring than either a short-term panel (such as those created in the special studies) or pure cross-sectional variation can capture.

In discussing the costing implications of use of the 3999 Model I have suggested costing procedures that can be used for accountables and collection volume, despite the fact that the Form 3999 dataset in its current form lacks volume measures for these mail streams. In an ideal world it would be desirable to correct this deficiency.

For the reasons described above, the highest priority should be placed on improvement of the parcel volume data collected by the Postal Service. The parcel volume measure currently contained in the national Form 3999 dataset, as described above, appears to have a strong relationship to deviation parcel volumes. However, that relationship is imperfect, and it is not clear what purpose, if any, the DOIS parcel measure serves in the Postal Service's carrier data systems. The Postal Service has itself suggested that there is little reliability or usefulness to the DOIS parcel counts (which are ultimately those provided in the nationwide Form 3999 dataset).

The improvement I am proposing should be an achievable goal. The Postal Service has invested heavily in scanners and other information technology to improve its ability to track the movement of parcels through its system. The existence of these systems and capabilities suggests that either now or in the near future the necessarily volume information will be available in an automated form to the Postal Service. Given the rapid growth in parcel volumes that has taken place, and the frequent statements by Postal Service management regarding the extent to which parcels will be the mainstay of its future business, the critical importance of improving the quality of parcel count data should be self-evident.[57]

The availability of improved volume count information will improve the accuracy of the costing parameters derived from the national Form 3999 dataset, and support the goal of moving away from costly special studies and establishing instead a system that can be updated annually and kept abreast of rapidly changing trends in the delivery environment.

[57] Note that I am not necessarily suggesting that DOIS data, which to my understanding is collected and recorded on a daily basis, needs to reflect accurate counts of deviation parcels and in-receptacle parcels. I recognize that daily collection of accurate parcel data would likely entail higher costs than providing the same on the periodic basis entailed in the Form 3999 evaluations. Even if the routine capture and reporting of parcel scanner date proves to be an unreachable near term goal, it does not appear overly burdensome to require the Postal Service once every three of four years when a city carrier route is being evaluated to conduct an accurate count of the number of parcels being delivered that day.